

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2003, 03:59:40 ; Search time 3019 Seconds  
(without alignments)  
8878.332 Million cell updates/sec

Title: US-09-735-713A-1  
Perfect score: 921  
Sequence: 1 atgagtctcaaatgcttat.....acatccaaactggtactaa 921

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hgt:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pt:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	920.2	99.9	921	6	AX180293	Sequence
2	920.2	99.9	1262	6	AX342644	Sequence
3	920.2	99.9	1568	6	AX180299	Sequence
4	908.2	98.6	909	6	AX180295	Sequence
5	771.2	83.7	1671	6	AX360091	Sequence
6	495	53.7	495	6	AX180297	Sequence
7	213.2	23.1	3689	5	XL081291	U81291 Xenopus lae
8	202.8	22.0	164732	9	AC104237	AC104237 Homo sapi
9	202.8	22.0	180707	2	AC012228	AC012228 Homo sapi
10	176.6	19.2	3028	5	AB070367	AB070367 Bufo japo
11	163.2	17.7	2627	6	AX480940	AX480940 Sequence
12	157	17.0	4628	5	XL081290	U81290 Xenopus lae
13	156.2	17.0	169388	2	AC129620	AC129620 Rattus no
14	114.6	12.4	2913	6	AX360086	Sequence
15	89.4	9.7	169388	2	AC129620	AC129620 Rattus no
16	83.6	9.1	2025	10	AF356627	AF356627 Mus muscu
17	80.4	8.7	2231	10	BC019485	BC019485 Mus muscu
18	76.2	8.3	2409	6	AX360098	Sequence
19	74.6	8.1	2672	6	AX207905	Sequence
20	74.6	8.1	3104	6	AX207903	Sequence
21	72.2	7.8	1850	5	AB038496	AB038496 Xenopus l
22	71	7.7	891	4	BOVERGS	L19663 Bos taurus
23	71	7.7	2581	6	I64576	I64576 Sequence 1
24	71	7.7	3860	4	BTU09859	U09859 Bos taurus
25	69.4	7.5	3586	4	PIGENPR	D30799 Sus scrofa
26	68.8	7.5	708	6	AX207901	Sequence
27	68.8	7.5	853	3	AF130842	Rhyzopert
28	68.8	7.5	3534	3	CIN431686	AF130842 Rhyzopert
29	68.4	7.4	2594	3	AY119618	AY119618 Drosophil
30	68.4	7.4	2752	10	MMU300738	AY119618 Drosophil
31	68.4	7.4	3749	3	DROSTUBBLE	AJ300738 Mus muscu
32	65.8	7.1	942	3	SSU79521	L11451 Drosophila
33	65.6	7.1	3696	6	AX207959	U79521 Scolopendra
34	65.6	7.1	3696	9	HSU09860	AX207959 Sequence
35	65.2	7.1	2331	4	AF395821	U09860 Human enter
36	64.6	7.0	2407	3	AY052044	AF395821 Oryctolag
37	63.2	6.9	1916	9	AF045649	AY052044 Drosophil
38	63.2	6.9	2087	9	HUMFXI	AF045649 Homo sapi
39	62.4	6.8	829	3	AF487426	M1142 Human facto
40	62.2	6.8	3183	10	BC029645	AF487426 Aedes aeg
41	60.8	6.6	891	3	AF508783	BC029645 Mus muscu
42	60.2	6.5	1303	10	AB032200	AF508783 Aedes aeg
43	60	6.5	2230	4	AB022425	AB032200 Mus muscu
44	59.8	6.5	3652	3	DMU18130	AB022425 Sus scrof
45	59	6.4	813	4	P1GF1XA	U18130 Drosophila
						M26235 Pig factor

ALIGNMENTS

RESULT 1  
AX180293  
LOCUS AX180293 921 bp DNA linear PAT 06-AUG-2001  
DEFINITION Sequence 1 from Patent WO0146407.  
ACCESSION AX180293  
VERSION AX180293.1 GI:15132262  
KEYWORDS human.  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 921)  
AUTHORS Walke,D.W., Turner,C.A., Abuin,A., Friedrich,G., Zambrowicz,B. and Sands,A.T.  
TITLE Polynucleotides encoding human protease homologs

JOURNAL Patent: WO 0146407-A 1 28-JUN-2001;

FEATURES Lexicon Genetics Incorporated (US)

Location/Qualifiers

source 1..921

/organism="Homo sapiens"

/db\_xref="taxon:9606"

BASE COUNT 242 a 195 c 240 g 242 t 2 others

ORIGIN

Query Match 99.9%; Score 920.2; DB 6; Length 921;

Best Local Similarity 100.0%; Pred. No. 8.6e-277;

Matches 921; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGTCTCAAAATGCTTATAAGCAGGAACAAGCTGATTTTACTACTAGGAATAGTCTTT 60

Db 1 ATGAGTCTCAAAATGCTTATAAGCAGGAACAAGCTGATTTTACTACTAGGAATAGTCTTT 60

Qy 61 TTGGAACRAGGTAATCTGCARCTCTTTTCGCTCCCAAGCTCCAGTCTGGGCAGAGT 120

Db 61 TTGGAACRAGGTAATCTGCARCTCTTTTCGCTCCCAAGCTCCAGTCTGGGCAGAGT 120

Qy 121 CTGTTTAAGGTACAGCCTTGGAAATTTTAAACATTTTTCAGTCGCATCTTGGAGGAGC 180

Db 121 CTGTTTAAGGTACAGCCTTGGAAATTTTAAACATTTTTCAGTCGCATCTTGGAGGAGC 180

Qy 181 CAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAGGCAAGAGCATATT 240

Db 181 CAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAGGCAAGAGCATATT 240

Qy 241 TGTGGAGGAAGCAATCTCAACACAGTGGGTGATCAGCGGGCTCACTGCATTGCAAAAC 300

Db 241 TGTGGAGGAAGCAATCTCAACACAGTGGGTGATCAGCGGGCTCACTGCATTGCAAAAC 300

Qy 301 AGAACATTTGCTACTTTGAATGTTTACTGCTGGAGAGTATGACITTAAGCCAGAGAC 360

Db 301 AGAACATTTGCTACTTTGAATGTTTACTGCTGGAGAGTATGACITTAAGCCAGAGAC 360

Qy 361 CCAGGAGAGCAAACTCTCACTATTGAAACTGTCTATCATATCATCCACATTTCTCCACCAAG 420

Db 361 CCAGGAGAGCAAACTCTCACTATTGAAACTGTCTATCATATCATCCACATTTCTCCACCAAG 420

Qy 421 AAACCAATGAGCATATGATATGCGCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCAC 480

Db 421 AAACCAATGAGCATATGATATGCGCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCAC 480

Qy 481 TTTGTGGGCCCCATATGCTCTCCAGAGCTCGGGAGCAATTTGAGGCTGGCTTTTATTGT 540

Db 481 TTTGTGGGCCCCATATGCTCTCCAGAGCTCGGGAGCAATTTGAGGCTGGCTTTTATTGT 540

Qy 541 ACAACTCGAGGCTGGGCGCGCTTAACTGAAGTGGCGTCTCTCACAAAGTCTTTCAGGAA 600

Db 541 ACAACTCGAGGCTGGGCGCGCTTAACTGAAGTGGCGTCTCTCACAAAGTCTTTCAGGAA 600

Qy 601 GTGAATCTGCTATTATTGACCTGGGAAGAGTGTGTGGCAGCTCTGTAAACACTAAAGAGG 660

Db 601 GTGAATCTGCTATTATTGACCTGGGAAGAGTGTGTGGCAGCTCTGTAAACACTAAAGAGG 660

Qy 661 CCCATCAGTGGGAAGACCTTTCTTTGCACAGGTTTCTCTGATGGAGGAGAGCAGCATGT 720

Db 661 CCCATCAGTGGGAAGACCTTTCTTTGCACAGGTTTCTCTGATGGAGGAGAGCAGCATGT 720

Qy 721 CAGGAGATTTCAGAGGTTTCACTCATGTGCGGAATAAGAAAGGGGCTTGACTCTGGCT 780

Db 721 CAGGAGATTTCAGAGGTTTCACTCATGTGCGGAATAAGAAAGGGGCTTGACTCTGGCT 780

Qy 781 GGTGTGACTTCTCGGGGTTTGGGCTGTGTGTCAGGCTGGAGAAACAATGTGAGGAAAAGT 840

Db 781 GGTGTGACTTCTCGGGGTTTGGGCTGTGTGTCAGGCTGGAGAAACAATGTGAGGAAAAGT 840

Qy 841 GATCAAGATCCCTCGGATCTTCACAGACATTAGTAAAGTCTTCTCGATCCAGGAA 900

Db 841 GATCAAGATCCCTCGGATCTTCACAGACATTAGTAAAGTCTTCTCGATCCAGGAA 900

Qy 901 CACATCCAAACTGGTAACATA 921

Db 901 CACATCCAAACTGGTAACATA 921

RESULT 2

AX342644

LOCUS

Sequence 41 from Patent WO0198468.

DEFINITION

AX342644

ACCESSION

AX342644.1

VERSION

AX342644.1

KEYWORDS

human.

ORGANISM

Homo sapiens

REFERENCE

1

Yue, H., Elliott, V.S., Gandhi, A.R., Lal, P., Au-Young, J.,

Tribouley, C.M., Deleage, A.M., Baughn, M.R., Nguyen, D.B., Lee, E.A.,

Hafalia, A., Khan, F.A., Wallia, N.K., Yao, M.G., Lu, D.A., Patterson, C.,

Tang, Y.T., Walsh, R.T., Azimzai, Y., Ramkumar, J., Xu, Y. and Reddy, R.

Patent: WO 0198468-A 41 27-DEC-2001;

Incyte Genomics, Inc. (US)

FEATURES

Location/Qualifiers

1..1262

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/note="Incyte ID No: 433459CB1"

BASE COUNT 354 a 265 c 306 g 337 t

ORIGIN

Query Match 99.9%; Score 920.2; DB 6; Length 1262;

Best Local Similarity 99.8%; Pred. No. 8.9e-277;

Matches 919; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGTCTCAAAATGCTTATAAGCAGGAACAAGCTGATTTTACTACTAGGAATAGTCTTT 60

Db 165 ATGAGTCTCAAAATGCTTATAAGCAGGAACAAGCTGATTTTACTACTAGGAATAGTCTTT 224

Qy 61 TTTGAACRAGGTAATCTGCARCTCTTTTCGCTCCCAAGCTCCAGTCTGGGCAGAGT 120

Db 225 TTTGAACRAGGTAATCTGCARCTCTTTTCGCTCCCAAGCTCCAGTCTGGGCAGAGT 284

Qy 121 CTGTTTAAGGTACAGCCTTGGAAATTTTAAACATTTTTCAGTCGCATCTTGGAGGAGC 180

Db 285 CTGTTTAAGGTACAGCCTTGGAAATTTTAAACATTTTTCAGTCGCATCTTGGAGGAGC 344

Qy 181 CAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAGGCAAGAGCATATT 240

Db 345 CAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAGGCAAGAGCATATT 404

Qy 241 TGTGGAGGAAGCATCTCTCACCACAGTGGGTGATCAGCGGGCTCACTGCATTGCAAAAC 300

Db 405 TGTGGAGGAAGCATCTCTCACCACAGTGGGTGATCAGCGGGCTCACTGCATTGCAAAAC 464

Qy 301 AGAACATTTGCTACTTTGAATGTTTACTGCTGGAGAGTATGACTTTAAGCCAGAGAC 360

Db 465 AGAACATTTGCTACTTTGAATGTTTACTGCTGGAGAGTATGACTTTAAGCCAGAGAC 524

Qy 361 CCAGGAGAGCAAACTCTCACTATTGAAACTGTCTATCATATCATCCACATTTCTCCACCAAG 420

Db 525 CCAGGAGAGCAAACTCTCACTATTGAAACTGTCTATCATATCATCCACATTTCTCCACCAAG 584

Qy 421 AAACCAATGAGCATATGATATGCGCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCAC 480

Db 585 AAACCAATGAGCATATGATATGCGCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCAC 644

Qy 481 TTTGTGGGCCCCATATGCTCTCCAGAGCTCGGGAGCAATTTGAGGCTGGCTTTTATTGT 540

Db 645 TTTGTGGGCCCCATATGCTCTCCAGAGCTCGGGAGCAATTTGAGGCTGGCTTTTATTGT 704

Qy 541 ACAACTCGAGGCTGGGCGCGCTTAACTGAAGTGGGCTCTCTCACAAAGTCTTTCAGGAA 600

Db 541 ACAACTCGAGGCTGGGCGCGCTTAACTGAAGTGGGCTCTCTCACAAAGTCTTTCAGGAA 600

Db 705 ACAACTGACGGCTGGGGCCGCTTAACCTGAAGGTGGGTCCTCTCACAAAGTCTTGCAGAA 764  
QY 601 GTGAATCTGCTATTTTGACCTGGGAGAGTGTGGCAGCTCTGTAAACACTAAAGAGG 660  
Db 765 GTGAATCTGCTATTTTGACCTGGGAGAGTGTGGCAGCTCTGTAAACACTAAAGAGG 824  
QY 661 CCATCAGTGGGAGAGCTTTCTTTTGACAGAGTGTGGCAGCTCTGTAAACACTAAAGAGG 720  
Db 825 CCATCAGTGGGAGAGCTTTCTTTTGACAGAGTGTGGCAGCTCTGTAAACACTAAAGAGG 884  
QY 721 CAGGAGATTACAGAGTTTACATGTCGCGGAATAAGAAAGGGGCTTGGCTGTGGCT 780  
Db 885 CAGGAGATTACAGAGTTTACATGTCGCGGAATAAGAAAGGGGCTTGGCTGTGGCT 944  
QY 781 GGTGTGATCTTCCCTGGGCTTTGGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 840  
Db 945 GGTGTGATCTTCCCTGGGCTTTGGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 1004  
QY 841 GATCAAGGATCCCTGGGATCTTACAGACATTAGTAAAGTGTCTTCCCTGGATCCAGAA 900  
Db 1005 GATCAAGGATCCCTGGGATCTTACAGACATTAGTAAAGTGTCTTCCCTGGATCCAGAA 1064  
QY 901 CACATCCAAACTGGTAACTAA 921  
Db 1065 CACATCCAAACTGGTAACTAA 1085

RESULT 3  
AX180299  
LOCUS AX180299 1568 bp DNA linear PAT 06-AUG-2001  
DEFINITION Sequence 7 from Patent WO0146407.  
ACCESSION AX180299  
VERSION AX180299.1 GI:15132265  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1568)  
AUTHORS Walke,D.W., Turner,C.A., Abuin,A., Friedrich,G., Zambrowicz,B. and Sands,A.T.  
TITLE Polynucleotides encoding human protease homologs  
JOURNAL Patent: WO 0146407-A 7 28-JUN-2001;  
Lexicon Genetics Incorporated (US)  
FEATURES  
Location/Qualifiers  
source  
1..1568  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 452 a 338 c 353 g 423 t 2 others  
ORIGIN

Query Match 99.9%; Score 920.2; DB 6; Length 1568;  
Best Local Similarity 100.0%; Pred. No. 9.1e-277;  
Matches 921; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGTCTCAAAATGCTTATAAGCAGGAAAGCTGATTTTACTACTAGGAATAGTCTTT 60  
Db 245 ATGAGTCTCAAAATGCTTATAAGCAGGAAAGCTGATTTTACTACTAGGAATAGTCTTT 304  
QY 61 TTTGAACRAGGTAATCTGCARCTTTTCGCTCCCAAGCTCCCAAGTGTGGGAGAGT 120  
Db 305 TTTGAACRAGGTAATCTGCARCTTTTCGCTCCCAAGCTCCCAAGTGTGGGAGAGT 364  
QY 121 CTGGTTAAGTACAGCTTGAATTTTAAACATTTTCACTTCAGTCGCATTTTGGAGGAGC 180  
Db 365 CTGGTTAAGTACAGCTTGAATTTTAAACATTTTCACTTCAGTCGCATTTTGGAGGAGC 424  
QY 181 CAAGTGGAGAGGTTTCTATCCCTGGCAGGTATCTCTGAAACAAAGGCAAGACATATT 240  
Db 425 CAAGTGGAGAGGTTTCTATCCCTGGCAGGTATCTCTGAAACAAAGGCAAGACATATT 484  
QY 241 TGTGGAGGAAGCATCGCTCTCACCACAGTGGGTGATCAGCGCGCTCAGTGCATTGCAAAAC 300  
|||||

Db 485 TGTGGAGGAAGCATCGCTCTCACCACAGTGGGTGATCAGCGCGCTCAGTGCATTGCAAAAC 544  
QY 301 AGAAACATGCTGCTACTTTGAATGTTACTGCTGGAGAGTATGACTTAAGCCAGACAGAC 360  
Db 545 AGAAACATGCTGCTACTTTGAATGTTACTGCTGGAGAGTATGACTTAAGCCAGACAGAC 604  
QY 361 CCAGGAGAGCAAACTCTCAGTATTTGAAACTGTTCATCATACATCCACATTTCTCCACCAAG 420  
Db 605 CCAGGAGAGCAAACTCTCAGTATTTGAAACTGTTCATCATACATCCACATTTCTCCACCAAG 664  
QY 421 AAACCAATGGACATATGATATTTGCCCTTTTGAAGATGGCTGGAGCCCTTCCAATTTGGCCAC 480  
Db 665 AAACCAATGGACATATGATATTTGCCCTTTTGAAGATGGCTGGAGCCCTTCCAATTTGGCCAC 724  
QY 481 TTTGTGGGGCCCATATGCTCTTCCAGAGCTGCGGAGCAATTTGAGGCTGTTTTATTGT 540  
Db 725 TTTGTGGGGCCCATATGCTCTTCCAGAGCTGCGGAGCAATTTGAGGCTGTTTTATTGT 784  
QY 541 ACAACTGCAGGCTGGGGCCGCTTAACTGAAGTGGGCTCCTCTCACAAAGTCTTTCAGAGAA 600  
Db 785 ACAACTGCAGGCTGGGGCCGCTTAACTGAAGTGGGCTCCTCTCACAAAGTCTTTCAGAGAA 844  
QY 601 GTCAATCTGCCATATTTTGACCTGGGAGAGTGTGGCAGCTCTGTAAACACTAAAGAGG 660  
Db 845 GTGAATCTGCCATATTTTGACCTGGGAGAGTGTGGCAGCTCTGTAAACACTAAAGAGG 904  
QY 661 CCATCAGTGGGAGAGCTTTCTTTTGCACAGTGTTCCTCTGATGGAGGAGAGCAGCATGT 720  
Db 905 CCATCAGTGGGAGAGCTTTCTTTTGCACAGTGTTCCTCTGATGGAGGAGAGCAGCATGT 964  
QY 721 CAGGAGATTACAGAGTTCACATCATGTGCGGAATAAGAAAGGGGCTTGGCTGTGGCT 780  
Db 965 CAGGAGATTACAGAGTTCACATCATGTGCGGAATAAGAAAGGGGCTTGGCTGTGGCT 1024  
QY 781 GGTGTGACTTCTCGGGGTTTGGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 840  
Db 1025 GGTGTGACTTCTCGGGGTTTGGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 1084  
QY 841 GATCAAGGATCCCTGGGATCTTACAGACATTAGTAAAGTGTCTTCCCTGGATCCAGAA 900  
Db 1085 GATCAAGGATCCCTGGGATCTTACAGACATTAGTAAAGTGTCTTCCCTGGATCCAGAA 1144  
QY 901 CACATCCAAACTGGTAACTAA 921  
Db 1145 CACATCCAAACTGGTAACTAA 1165

RESULT 4  
AX180295  
LOCUS AX180295 909 bp DNA linear PAT 06-AUG-2001  
DEFINITION Sequence 3 from Patent WO0146407.  
ACCESSION AX180295  
VERSION AX180295.1 GI:15132263  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 909)  
AUTHORS Walke,D.W., Turner,C.A., Abuin,A., Friedrich,G., Zambrowicz,B. and Sands,A.T.  
TITLE Polynucleotides encoding human protease homologs  
JOURNAL Patent: WO 0146407-A 3 28-JUN-2001;  
Lexicon Genetics Incorporated (US)  
FEATURES  
Location/Qualifiers  
source  
1..909  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 237 a 193 c 238 g 239 t 2 others  
ORIGIN

Query Match 98.6%; Score 908.2; DB 6; Length 909;  
Best Local Similarity 100.0%; Pred. No. 4.9e-273;







```
repeat_region 25237. 25470
                /rpt_family="L2"
repeat_region complement(25810. 25919)
                /rpt_family="MIR"
repeat_region 28709. 29154
                /rpt_family="LTR16A1"
repeat_region 29252. 29284
                /rpt_family="(CAAAAA)n"
repeat_region 29612. 30220
                /rpt_family="L1"
```

```
repeat_region 30485 30788
/rpt_family="L1"
30485 30788
```

```
repeat_region
/rpt_family="L1"
30486. .30788
/rpt_family="AluX"
repeat_region
complement(30809. .31550)
/rpt_family="L1p"
repeat_region
31551. .32627
/rpt_family="L1PA3"
repeat_region
complement(32836. .33451)
/rpt_family="L1p"
```

repeat_region	complement(342836, .35431)
	/rpt_family="L1P"
repeat_region	complement(33489, .33965)
	/rpt_family="L1"
repeat_region	33977, .34152
	/rpt_family="L1"
repeat_region	complement(34157, .34342)
	/rpt_family="MER46A"
repeat_region	34343, .35088
	/rpt_family="L1MDA"
repeat_region	35089, .35303
	/rpt_family="L2"
repeat_region	35968, .36617
	/rpt_family="L1PA13"

```

unsure          0: Gaps      0:
/note="single clone coverage"
    very Match      22.0%; Score 202.8; DB 9; Length 164732;
    distLocal Similarity 92.68; Pred. No. 9.9e-52;
    -ches 213; Conservative 0; Mismatches 17; Indels 0: Gaps 0:

692 GTTTTCCTCATGGAGGAGACGCATCTCAGGAGATTTCAGGAGGTTCACCTCATGTGCC 751
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
51264 GTGATGCTAATAAGAGCCCTTTTCCCTCTCGCAGGAGATTTCAGGAGGTTCACCTCATGTGCC 51323

752 GGAAATAAGAAAGGGCCCTGGACTCTGGCTGTGTGTGACTTCCTCTGGGTTTGGGCTGTGGTC 811
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
51324 GGAATAAGAAAGGGCCCTGGACTCTGGCTGTGTGACTTCCTCTGGGTTTGGGCTGTGGTC 51383

812 GAGGCTGCAGAAACAATCTGAGGAAAACGTGATCAAGGATCCCCTGGATCTTCACAGACA 871
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
51384 GAGGCTGCAGAAACAATCTGAGGAAAACGTGATCAAGGATCCCCTGGATCTTCACAGACA 51443

872 TTAGTAAGAAGTGCTTTCCTGGATCCACGAAACACATCCCAAACCTGGTAACATA 921
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
51444 TTAGTAAGAAGTGCTTTCCTGGATCCACGAAACACATCCCAAACCTGGTAACATA 51493

T 9
S 5
C 22228/C
AC012228 180707 bp DNA linear HTG 17-AUG-2002
Homo sapiens chromosome 11 clone RP11-439A13 map 11, *** SEQUENCING
IN PROGRESS ***, 2 ordered pieces.
AC012228
AC012228.8 GI:22123492
HTG: HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 180707)
Birren,B., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 11, clone RP11-439A13
Unpublished
2 (bases 1 to 180707)

```

```

every Match      22.0%; Score 202.8; DB 9; Length 164732;
1st Local Similarity 92.6%; Pred. No. 9.9e-52;
Mismatches 213; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

```

692 GTTTCCTGATGGAGGGAGACGCAGTCTCAGGAGATTCAGGAGGTTCACTCATGTGCC 751  
61264 CTGATGCTTAATAGAGGCTTTTCCCTCTCGCAGGAGATTCAGGAGGTTCACTCATGTGCC 51323			
752 GGAATTAAGAAAGGGGCCCTGGACTCTGGCTGGTGTGACTCTCCCTGGGTTTGGGCTGTGGTC 811			
61324 GGAATTAAGAAAGGGGCCCTGGACTCTGGCTGGTGTGACTCTCCCTGGGTTTGGGCTGTGGTC 51383			
812 GAGGCTGGAGAAACAATCTGAGGAAAAGTGATCAAGGATCCCTCGGATCTTCACAGACA 871			
61384 GAGGCTGGAGAAACAATCTGAGGAAAAGTGATCAAGGATCCCTCGGATCTTCACAGACA 51443			
872 TTAGTAAAGTGGTTTCCCTGGATCCACGAACACATCCCAACCTGGTAACTAA 921			
61444 TTAGTAAAGTGGTTTCCCTGGATCCACGAACACATCCCAACCTGGTAACTAA 51493

AC012228 180707 bp DNA linear HTG 17-AUG-2002  
 Homo sapiens chromosome 11 clone RP11-439A13 map 11, \*\*\* SEQUENCING  
 IN PROGRESS \*\*\*, 2 ordered pieces.  
 AC012228  
 AC012228.8 GI:21213492  
 HTG; HTGS\_PHASE2; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.  
 human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 180707)  
 Birren,B., Nusbaum,C. and Lander,E.  
 Homo sapiens chromosome 11, clone RP11-439A13  
 Unpublished  
 2 (bases 1 to 180707)

## AUTHORS

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckerly,R., Boguslavkiy,L., Boukhgalter,B., Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A., Doyle,M., Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Gage,D., Ferrerira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J., Gadyana,S., Grant,G., Hagos,B., Hearford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,J., Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

## TITLE

Submitted (21-OCT-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

## REFERENCE

## AUTHORS

3 (bases 1 to 180707)  
 Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gadyana,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,I., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

## TITLE

## JOURNAL

Submitted (17-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Aug 6, 2002 this sequence version replaced gi:21362200.  
 All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L2182

Center clone name: 439\_A\_13

-----

\* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

\* This sequence will be replaced

\* by the finished sequence as soon as it is available and

\* the accession number will be preserved.

\* 1 71362: contig of 71362 bp in length

\* 71363 71462: gap of 100 bp

\* 71463 180707: contig of 109245 bp in length.

Location/Qualifiers

1. .180707

/organism="Homo sapiens"

/db\_xref="taxon.9606"

/chromosome="11"

/map="11"

/clone="RP11-439A13"

/clone\_11b="RPC1-11 Human Male BAC"

BASE COUNT 47741 a 39327 c 40651 g 52888 t 100 others

## ORIGIN

Query Match 22.0%; Score 202.8; DB 2; Length 180707;  
 Best Local Similarity 92.6%; Pred. No. 1e-51;  
 Matches 213; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
 QY 692 GTTFTTCTCTGAGGAGACGCATGTCTCAGGAGATTCAGGAGGTTCACATCATGTGCC 751  
 Db 165369 GTGATGCTAATAGAGGCTTTTCCTCTGCAGGAGATTCAGGAGGTTCACATCATGTGCC 165310  
 QY 752 GGAATAAGAAAGGGCTGGACTCTGGCTGGTGTGACTTCTCTGGGGTTTGGGCTGTGGTC 811  
 Db 165309 GGAATAAGAAAGGGCTGGACTCTGGCTGGTGTGACTTCTCTGGGGTTTGGGCTGTGGTC 165250  
 QY 812 GAGCTCGAGAAACAATGTGAGGAAAGTATCAAGGATCCCTGGGATCTTCACAGACA 871  
 Db 165249 GAGCTCGAGAAACAATGTGAGGAAAGTATCAAGGATCCCTGGGATCTTCACAGACA 165190  
 QY 872 TTAGTAAAGTCTTCTCTGGATCCACGAACACATCCAAACTGGTAACTAA 921  
 Db 165189 TTAGTAAAGTCTTCTCTGGATCCACGAACACATCCAAACTGGTAACTAA 165140

## RESULT 10

## AB070367

## LOCUS

AB070367 3028 bp mRNA linear VRT 14-MAR-2002  
 Bufo japonicus mRNA for oviductin, complete cds.

## ACCESSION

## AB070367

## VERSION

## AB070367.1

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## FEATURES

## source

## CDS

## FEATURES

## Location/Qualifiers

## 1. .3028

## /organism="Bufo japonicus"

## /db\_xref="taxon.8387"

## /sex="female"

## /tissue\_type="oviductal pars recta"

## 85. .3009

## /function="cleaves egg envelope glycoproteins"

## /note="oviduct serine protease"

## /codon\_start=1

## /product="oviductin"

## /protein\_id="BAB63372.1"

## /db\_xref="GI:15277254"

## /translation="MAETSIFFIMLTVMIGVGRVTDSPGRVSRCCRPAANTSVSY

## GLLSRVVGGTSVAVKSPWMLKRDGKHFCGGTIIISDKVLTAAHCVLKNFEFOVS

## VSIQDHDFAVYERSEORFAIKSVKHPNFKSPFNLDLAILVELVSTITDKDQAPAC

## LPSQDDVFPTGTLTCLMALGRLGLOENGLDSSLOKVLPLLEYRCLSIMETVDRLLAF

## ETVVCAGFFEGGKACQDSSGFFLCORSQGRVWLVGVTSWGLGCKARKWADNILDPE

## SKSGPGVFTDIIQLLNLWLSNLNDKDPDPFQVQCSQDNDIEKRTGTEILLPTGYKK

## YYSNNEKCIWTIIVPRGKHILLTFKFNVECDYSCDLDLVITYSALGRLLIGLFCGDS

## PRPLLADASITLKFISDFEYTKFSLEFVPEPDTYQDSCGSAVIFEEGEIQTM

## NHPLYSSHANQWVHSVPANYIIKIFTLFVEVEPSEGCIFDRLVYHDLQGTWVAGF

FCGFLPDPVLSVSNMQITFTSDYSANYLGFRAVIVSFVLPSSPVKPEKGNQPRKNO  
DAMQHFEGCVDPRFLYHNLKAEAEAMNSPWHVSINFGNKHVCNGAILSKTF  
VWTSANCAVDREPERPSICLIVAGLHDLESSINTOKRVEYVIVHPDNRLSKDLDVAL  
IHVORPQNSYVOPICLPDGHSLRLEFSKLCVSGWDLNVELSLKQOLEVPGVLMDDV  
CKKYDGLTDMFCWAPAEEDNASCLAQSGAPLVQCSAPGYAIFGIVSRVGCNET  
PKAGYSSVLEFIPWIMETILSVAGIIDTSEPHPLFPDPKPSQOKALLPDSPPSS  
SODIYVTKDVLSPQELKVASGDGPGEGRCQLIFQAPGHEFTLLTFKQLSHEH  
YSLIIEGASSNKTFFKAQLMBEIKPTIMKSAGAVITILEASSTAQDSALHLWLSYSFHN  
QN"

misc\_feature 232..990  
/note="encodes protease domain"  
misc\_feature 1018..1359  
/note="encodes CUB domain"  
misc\_feature 1387..1725  
/note="encodes CUB domain"  
BASE COUNT 800 a 696 c 727 g 805 t  
ORIGIN

Query Match 19.2%; Score 176.6; DB 5; Length 3028;  
Best Local Similarity 53.1%; Pred. No. 1e-43;  
Matches 402; Conservative 0; Mismatches 349; Indels 6; Gaps 1;  
Qy 158 TCAGTCGCATCTTGGAGGAAGCCAAAGTGGAGAGGGTTCCTATTCCTGGCAGGTATCTC 217  
Db 224 TCTCTCGTATTGTTGGAGGGACATCTGCAGTCAAGGGAAGCCCATGGATGGTGTCTG 283  
Qy 218 TCAACAAAGGCAGAGCATATTTCTGGAGGAAGCATCTCTCACACAGTGGGTGATCA 277  
Db 284 TAAAAGAGATGGGAAGCACCTCTCTGGAGGAACCATTTATTAGCGATAAATACGTGCTA 343  
Qy 278 CGGGGCTCACTGCTATTCGAACAGAAACATTTGCTCTACTTTTGAATGTTACTGCTGGAG 337  
Db 344 CGGCTGCGCAUATGCTCTTAGAAAAAATCTTGAATTTCAAGTGAGCGCTTCGATTGGAG 403  
Qy 338 AGTATGACTTAAAGCAGACAGCCAGGAGAGCAAACTCTCACTATTGAAACTGTCATCA 397  
Db 404 ATCAGATTTCGCTGTATATGAAGAGAGTGAACACGCTTTCGCCATAAAATCTGTTTCA 463  
Qy 398 TACATCCACATTTCCACCAAGAAACCAATGGACTATGATATTGCCCTTTTGAAGATGG 457  
Db 464 AGCACCCCACTTCAACCGAGCAGGCCCTTCAACTATGACTCGCCATCTGAGCTGG 523  
Qy 458 CTGGAGCCTTCCAATTTGGCCACTTTGTGGGGCCCATATGTCTCCAGAGCTGCGGGAGC 517  
Db 524 TGGAAAGCATCACATTTGATAAAGATATCCAACCGCATGCTACCGAGCCCGATGATG 583  
Qy 518 AATTGAGCGCTGTTTATTGTTACAACTGCAGGCTGGGCGGCTTAAGTGAAGTGGCG 577  
Db 584 TATTCCAAAGAACTCTGTATGGCCTTTGTGGGAGCTGTGGTGGGAGCTGCGAGGAAATGAA 643  
Qy 578 TCCTCTCACAAGTCTTGAGGAAGTGAATCTGCCTATTATTTGACCTTGGGAAGAGTGTGG 637  
Db 644 GTTACCTAGCAGCTGCAAAAAGTTGTTTACCTCTCATTTAGTACAGAAAGTGTCTGA 703  
Qy 638 CAGCTCTGTTAACTAAAGAGGCCCATCAGTGGGAACACCTTCTTTGACAGGTTTTC 697  
Db 704 GCATCATGGAGACTGTGATGCGGCTTTGGCATTTTGAGACTGTGGTTTGTGCAAGGATTC 763  
Qy 698 CTGATGGAGGAGAGACCATGTTCAGGAGATTCAGGAGGTTCACTCATGTGCCGGAATA 757  
Db 764 CAGAAGGGGTGAAGACCGGTGTACAGGTGACTCTGGAGGTCTTCTCTGCGCCAGAGA 823  
Qy 758 AGAAGGGGCTGACTCTGGCTGTGACTTCTCTGGGGTTTGGGCTGTGGTCTGAGCT 817  
Db 824 GCCAAGGAAGTGGGTTTGGTGTGTTAACTCTCTGGGGATTAGGATGTGCACGAAAT 883  
Qy 818 GGAGAAACATGTGAGGAAAAGTG-----ATCAAGGATCCCTGGGATCTTTCACAGACA 871  
Db 884 GGGCAGACAATATCTGGACCCAGTGGAAAGCAAAAGTTTACCAGGTCTTTTTCACGGATA 943  
Qy 872 TTAGTAAGTCTTCTTCGGATCCACGACACATCCA 908  
Db 944 TTCAGAGGCTTCTCAACTGGCTCAGTGAACACCTCAA 980

RESULT 11  
AX480940  
LOCUS AX480940  
DEFINITION Sequence 36 from Patent WO0246383.  
ACCESSION AX480940  
VERSION AX480940.1 GI:22217579  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS  
Yue, H., Azimzai, Y., Kallick, D. A., Baughn, M. R., Griffin, J. A.,  
Swarnakar, A., Lal, P. G., Walla, N. K., Hafalia, A. J., Gandhi, A. R.,  
Au-Young, J., Elliott, V. S., Ramkumar, J., Thangavelu, K., Lu, Y.,  
Warren, B. A., Lee, E. A., Tribouley, C. M., Arvizu, C.,  
Deleane, A. M., Yao, M. G., Khan, F. A. and Sanjanwala, M. M.  
Protein modification and maintenance molecules  
Patent: WO 0246383-A 36 13-JUN-2002;  
Incyte Genomics, Inc. (US)  
FEATURES  
Location/Qualifiers  
source  
1..2627  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 7484157CBI"  
BASE COUNT 739 a 561 c 636 g 691 t  
ORIGIN

Query Match 17.7%; Score 163.2; DB 6; Length 2627;  
Best Local Similarity 51.7%; Pred. No. 1.6e-39;  
Matches 400; Conservative 0; Mismatches 368; Indels 6; Gaps 1;  
Qy 148 TTTAATATTTTCAGTCGCATCTTGGAGGAAGCCAAAGTGGAGAGGGTTCCTATCCCTGG 207  
Db 156 TCTAGATTTCTCTCTAGAAATTAGTATTGGAGAAATTCAAACAGTCACTGGCATCCATGG 215  
Qy 208 CAGGTATCTCTGAACAAGGCAGAGCATATTTGTGGAGGAGCATGCTCTCACACAG 267  
Db 216 CAGGTCTCTCTTAAATCAGATGAGCACCCTCTCTGTGGAGGAGCTTGATTCAGAAAGT 275  
Qy 268 TGGTGATCAGCGCGCTCACTGCTTCAAAACAGA-----AACATTTGTGTCTACTTTG 321  
Db 276 CGGTTGTTTACAGCAGCACACTGCTGGACAGCCTCAGTGAGAACAGCTGAAGAATATA 335  
Qy 322 AATGTTACTGCTGGAGAGTATGACTTAAGCCAGACAGACCCAGGAGAGCAAACTCTCACT 381  
Db 336 ACTGTGACTTCTGGGGAGTACAGCCTCTTTCAGAAAGGATAAGCAAGACAGAAATATCCT 395  
Qy 382 ATTGAAATGTCATCATACATCCACATTTCTCCACCAAGAACCAATGGACTATGATATT 441  
Db 396 GTCTCAAAATATTACCCATCTCTGAATACACAGCCGTGAATATATGAGTCTCTGATTT 455  
Qy 442 GCCCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCACTTTTGGGGCCCATATGTCTT 501  
Db 456 GCAGTCTGTATCTAAAAACACAAAGTCAAGTTTGGAAATGCTGTTCAAGCAATCTGTCTT 515  
Qy 502 CCAGAGCTCGGGAGCAATTTGAGCTGGTATTATTTGTACAACTGCAGGCTGGGGCCGC 561  
Db 516 CTTGACAGCGATGATAAAGTTGAACCCAGGAATTTCTCTTATCCAGTGGATGGGGCAAG 575  
Qy 562 TTAAGTGAAGTGGCGTCTCTCAACAGTCTTTCAGAGGAAGTGAATCTGCCATTTTGTGACC 621  
Db 576 ATTTCCAAAGATCAGAATATTCAATGTCTACAAAGAAATGGAACTTCCCATCATGAT 635  
Qy 622 TGGGAAGATGTGTGGCAGCTCTGTAACTAAAGAGGCCCATCAGTGGGAAGACTTTT 681  
Db 636 GACAGAGCGTGTATCTGTCTCAAGAGCATGAACCTTCCCTCCCTCGGAGGAGCATG 695  
Qy 682 CTTTGCACAGGTTTTCTGTATGGAGGAGAGAGCATGTCAGGGAGATTCAGGAGGTTCA 741  
Db 696 CTGTGTCTGGCTTCCCTGATTGGGGAATGGACGCTGCCAGGGGGAGTCTCTGGAGGACCA 755



Db 536 AAGATTGAAGCGGACCTTGTGTGTTCCAGCGCTGGGGAAGGCTTGAAGAAAATGGG 595  
 QY 577 GTCTCTCACAAGTCTTGCAGGAAGTGAATCTGCCTATTTTTCACCTGGGAAAGAGTGTGTG 636  
 Db 596 GACTTGTCTCCCAAGTCTTACGAAGAAGTAAACATGCCAGTGTGATTAATGGCACCTGCCCAT 655  
 QY 637 GCAGCTCTCTTAACACTAAAGAGGCGCCATCAGTGGGAAGACCTTTCTTTTGCACACAGTGT 696  
 Db 656 GCTGTGTGTGGAGCCCATAGGCGCATCCGTTAGATGACACAATGTTATGTGCTGGATTTC 715  
 QY 697 CTTGATGGAGGAGAGACCCATCTCAGGAGAGATTACAGGAGTTCACTATGTCGCCGGAAT 756  
 Db 716 CCAGAAGGAGGATGATGATCCCTGCCAGGCGATTCTGGAGGACCCCTTTGTATGCAGACGG 775  
 QY 757 AAGAAAGGCGCTGCTGCTGTGTGACTTCTCTGGGTTTGGCTGTGTGCTCAGGC 816  
 Db 776 CGCTCTGGGGTGTGTTCTTGTGCTGGATGTATCTCTGGGGCTTGGCTGTGGGAGAAGC 835  
 QY 817 TGGAGAAACATGTTGAGGAAAAG 839  
 Db 836 TGGGGTGCACAAACAGATCATACG 858

## RESULT 13

AC129620

LOCUS

DEFINITION

Rattus norvegicus clone CH230-388116, \*\*\* SBUENCING IN PROGRESS

\*\*\*, 56 unordered pieces.

AC129620

AC129620.1 GI:22024303

HTG: HTGS\_PHASE1.

SOURCE

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 169388)

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
 Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,  
 Barbara,J., Benton,J., Blmage,K., Blankenburg,K., Bonnin,D.,  
 Bouck,J., Bowie,S., Brleva,M., Brown,M., Bryant,N.P.,  
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
 Carron,F.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,  
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,  
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,  
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
 Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,  
 Homsli,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,  
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
 Kratochvic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H.,  
 Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
 Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M.,  
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
 Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G.,  
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,I.,  
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
 Rivers,M., Rojas,A., RojuboKan,I., Rolfe,M., Ruiz,S., Savery,G.,  
 Scherfer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,  
 Sodergren,E., Sonake,T., Sparks,A., Stanley,H., Stone,H.,  
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
 Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,

Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
 Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,  
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
 Weinstein,G. and Gibbs,R.  
 Unpublished  
 2 (bases 1 to 169388)  
 Worley,K.C.  
 Direct Submission  
 Submitted (31-JUL-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 ----- Genome Center  
 Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 ----- Project Information  
 Center project name: KAHL  
 Center clone name: CH230-388116  
 ----- Summary Statistics  
 Sequencing vector: Plasmid:  
 Chemistry: Dye-terminator Big Dye: 100% of reads  
 Assembly program: Phrap: version 0.990329  
 Consensus quality: 136807 bases at least Q40  
 Consensus quality: 141381 bases at least Q30  
 Consensus quality: 145526 bases at least Q20  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 56 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 1002: contig of 1002 bp in length  
 \* 1003 1102: gap of unknown length  
 \* 1103 2132: contig of 1030 bp in length  
 \* 2133 2232: gap of unknown length  
 \* 2233 3358: contig of 1126 bp in length  
 \* 3359 3458: gap of unknown length  
 \* 3459 5023: contig of 1565 bp in length  
 \* 5024 5123: gap of unknown length  
 \* 5124 6663: contig of 1540 bp in length  
 \* 6664 6763: gap of unknown length  
 \* 6764 8065: contig of 1302 bp in length  
 \* 8066 8165: gap of unknown length  
 \* 8166 9312: contig of 1147 bp in length  
 \* 9313 9412: gap of unknown length  
 \* 9413 11980: contig of 2568 bp in length  
 \* 11981 12080: gap of unknown length  
 \* 12081 13407: contig of 1327 bp in length  
 \* 13408 13507: gap of unknown length  
 \* 13508 15457: contig of 1950 bp in length  
 \* 15458 15557: gap of unknown length  
 \* 15558 16567: contig of 1010 bp in length  
 \* 16567 16667: gap of unknown length  
 \* 16668 17971: contig of 1304 bp in length  
 \* 17972 18071: gap of unknown length  
 \* 18072 19740: contig of 1669 bp in length  
 \* 19741 19840: gap of unknown length  
 \* 19841 21942: contig of 2102 bp in length  
 \* 21943 22042: gap of unknown length  
 \* 22043 23410: contig of 1368 bp in length  
 \* 23411 23510: gap of unknown length  
 \* 23511 25146: contig of 1636 bp in length  
 \* 25147 25246: gap of unknown length  
 \* 25247 26357: contig of 1111 bp in length  
 \* 26358 26457: gap of unknown length  
 \* 26458 28550: contig of 2093 bp in length

TITLE  
JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

## COMMENT

* 151752	160340:	contig of 8589 bp in length
* 160341	160440:	gap of unknown length
* 160441	169388:	contig of 8948 bp in length.
FEATURES	Location/Qualifiers	
source	l..169388	
	/organism="Rattus norvegicus"	
Query Match	17.0%; Score 156.2; DB 2;	Length 169388;
Best Local Similarity	86.1%; Pred. No. 4e-37;	
Matches	173; Conservative	0; Mismatches 28; Indels 0; Gaps 0;
Qy 721	CAGGAGATTACAGAGGTTCACATCATGTGCCGGAAATAAGAAAGGGCCTTGACTCTGCCT	780
Db 98817	CAGGAGATTACAGAGGTTCACATCATGTGCCGGAAATAAGAAAGGGCCTTGACTCTGCCT	98876
Qy 781	GCTGTGATTTCTTCCCTGGGCTGTGGCGTGTGGCGAGGCTGGAGAACATGTGAGAAAAGT	840
Db 98877	GGAGTGATTTCTTCCCTGGGCTGTGGCGTGTGGCGAGGCTGGAGAACATGTGAGAAAAGT	98936
Qy 841	GATCAAGGATCCCCCTGGGATTTCCACAGACATTAAGTAAGTGTCTTCTGGATCCACGAA	900
Db 98937	GAACAGGATCTCTGGGATTTCCACAGATCTTAGGAGAGTGCTTCTGGATCCATGAA	98996
Qy 901	CACATCCAAACTGGTAACATAA 921	
Db 98997	CACGTCCAGACTGGTAATGAA 99017	
RESULT 14		
AX360086		
LOCUS	AX360086	2913 bp DNA linear PAT 13-FEB-2002
DEFINITION	Sequence 42 from Patent WO0200860.	
ACCESSION	AX360086	
VERSION	AX360086.1 GI:18675712	
KEYWORDS	human.	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 Plozman,G., Whyte,D., Sudarsanam,S., Manning,G., Caenepeel,S. and AUTHORS Charlyczak,G. TITLE Novel proteases JOURNAL Patent: WO 0200860-A 42 03-JAN-2002;	
FEATURES	Location/Qualifiers	
source	l..2913	
BASE COUNT	819 a 590 c 698 g 806 t	
ORIGIN		
Query Match	12.4%; Score 114.6; DB 6;	Length 2913;
Best Local Similarity	52.4%; Pred. No. 2.7e-24;	
Matches	252; Conservative	0; Mismatches 229; Indels 0; Gaps 0;
Qy 435	TGATATTGCCCTTTGAAGATGGCTGGAGCCTTCCAATTTGGCACATTTGTGGGGCCCAT	494
Db 9	TGATATTGCCCTTTGAAGATGGCTGGAGCCTTCCAATTTGGCACATTTGTGGGGCCCAT	68
Qy 495	ATGTCTTCCAGAGCTGGGGAGCAATTTGAGGCTGGTTTTATTATTGTACAACCTGCAGGCTG	554
Db 69	CTGTCTTCCAGAGCTGGGGAGCAATTTGAGGCTGGTTTTATTATTGTACAACCTGCAGGCTG	128
Qy 555	GGGGCGGCTTAACCTGAAGTGGCGTCTCTCACAAGTCTTGCGAGGAAGTGAATTCGCCTAT	614
Db 129	GGGCAAGATTTTCCAAAAACATCAGAATATTCAAATGTCTTACAAGAAATGGAACCTTCCCAT	188
Qy 615	TTTGACCTGGGAAGCTGTGTGGCAGCTCTTAACTTAACACTAAAGAGGCCCATCATCGTGGAA	674
Db 189	CATGGATGACAGACGGTGTAACTACTGTGCTCAAGAGCATGAACCTCCCTCCCTGGGAAG	248
Qy 675	GACCTTTCTTTGACACAGTTTTCTCTGATGGAGGAGAGAGCATGTGTCAGGAGATTCAGG	734





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2003, 04:03:10 ; Search time 275 seconds  
(without alignments)  
7542.147 Million cell updates/sec

Title: US-09-735-713A-1

Perfect score: 921

Sequence: 1 atgagtctcaaatgcttat.....acatccaactggtactaa 921

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002.\*

- 1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	920.2	99.9	921	AAD09328	Human protease hom
2	920.2	99.9	1262	ABK12903	Human protease PRT
3	920.2	99.9	1568	AAD09331	Human protease hom
4	908.2	98.6	909	AAD09329	Human protease hom
5	771.2	83.2	1671	ABK31790	DNA encoding novel
6	546.2	59.3	637	AAH98934	Toad EST-derived c
7	495	53.7	495	AAD09330	Human protease hom
8	260	28.2	324	AAC05816	Human secreted pro
9	163.2	17.7	2742	AAS88433	DNA encoding novel

c	10	163.2	17.7	3309	23	AAS87759	DNA encoding novel
	11	163.2	17.7	4729	23	AAS72057	DNA encoding novel
	12	163.2	17.7	4729	23	AAS74405	DNA encoding novel
	13	163.2	17.7	4729	23	AAS78775	DNA encoding novel
	14	163.2	17.7	4729	23	AAS84074	DNA encoding novel
	15	161.6	17.5	3096	24	ABQ86158	Novel human gene.
	16	114.6	12.4	2913	24	ABK31785	DNA encoding novel
	17	76.2	8.3	1327	20	AZ34034	Human EST DNA35597
	18	76.2	8.3	1327	21	AAC78495	Human EST DNA35597
	19	76.2	8.3	1327	21	AAC58237	Human EST (express
	20	76.2	8.3	2409	24	ABK31797	DNA encoding novel
	21	76.2	8.3	3143	20	AAC34033	Human PRO618 nucle
	22	76.2	8.3	3143	21	AAC78494	Human PRO618 (UNQ3
	23	76.2	8.3	3143	21	AAC58236	Human PRO618 nucle
	24	74.6	8.1	2672	22	AAD13117	Human membrane-ty
	25	74.6	8.1	3104	22	AAD13116	Human membrane-ty
	26	71	7.7	708	21	AAK99151	708nt DNA sequence
	27	71	7.7	708	24	ABK51438	DNA encoding cattl
	28	71	7.7	2581	15	AAQ70104	Bovine enterokinas
	29	68.8	7.5	708	22	AAD13115	Human membrane-ty
	30	68.4	7.4	3736	23	ABL28971	Drosophila melanog
	31	65.6	7.1	3696	22	AAD13165	Human enterokinase
	32	64.6	7.0	3196	23	ABL26453	Drosophila melanog
	33	63.2	6.9	2087	24	ABK83923	Human cDNA differe
	34	54.6	5.9	973	24	AAL45562	Lumbricus rubellus
	35	54.4	5.9	936	22	AAF75659	Human ztrypl degen
	36	54	5.9	796	24	ABK30337	Human G-protein-co
	37	53.2	5.8	734	24	ABK16738	Epithelin-like ser
	38	53.2	5.8	927	24	ABK16740	Epithelin-like ser
	39	53.2	5.8	2267	24	ABN95751	Gene #2249 used to
	40	53.2	5.8	2267	24	ABL67214	Thyroid cancer rel
	41	53.2	5.8	3387	24	ABK31781	DNA encoding novel
	42	53.2	5.8	3711	24	ABK86141	DNA encoding huma
	43	53	5.8	973	22	AAD093445	Human plasminogen
	44	53	5.8	1476	22	AAS64180	Human prostate cDN
	45	53	5.8	1476	22	AAH93944	p1000c open readin

#### ALIGNMENTS

RESULT 1  
AAD09328  
ID AAD09328 standard; cDNA; 921 BP.  
XX  
AC AAD09328;  
XX  
DT 10-SEP-2001 (first entry)  
XX  
DE Human protease homologue cDNA #1.  
XX  
KW Human; protease homologue; novel human protein; NHP; therapy;  
KW pharmacogenomic application; physiological disorder; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT CDS  
FT  
FT Location/Qualifiers  
FT 1..921  
FT /\*tag= a  
FT /product= "Human protease homologue #1"  
FT 67..69  
FT /\*tag= b  
FT /note= "Encodes Arg"  
FT 68  
FT /\*tag= c  
FT /note= "This degenerate base represents a  
FT polymorphic site"  
FT 82..84  
FT /\*tag= d  
FT /note= "Encodes Ala"  
FT 82  
FT /\*tag= e  
FT /note= "This degenerate base represents a

FT .. polymorphic site"  
PN WO200146407-A1.  
XX 28-JUN-2001.  
XX 12-DEC-2000; 2000WO-US33738.  
XX 23-DEC-1999; 99US-0171566.  
PR (LEXI-) LEXICON GENETICS INC.  
PA  
XX Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;  
XX WPI; 2001-408641/43.  
DR P-PSDB; AAE04733.  
XX  
XX Polynucleotide encoding novel human protease homologs, useful for  
PT identifying agonist, antagonist or modifiers or for producing  
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic  
PT applications -  
XX  
XX Claim 1; Page 28; 32pp; English.  
XX  
XX The present sequence is a cDNA encoding novel human protein (NHP),  
CC known as human protease homologue. NHP shares structural similarity  
CC with animal proteases, particularly trypsin-like protease such  
CC as oviductin, plasminogen activator and human plasma kallikrein  
CC precursor. NHP and its cDNA are useful in therapeutic, diagnostic and  
CC pharmacogenomic applications. NHP sequences are useful for identifying  
CC agonists, antagonists and modulators and also for producing antibodies  
CC useful in diagnosis, drug screening, clinical trial monitoring and in  
CC treatment of physiological disorders.  
XX  
XX Sequence 921 BP; 242 A; 195 C; 240 G; 242 T; 2 other;

Query Match 99.9%; Score 920.2; DB 22; Length 921;  
Best Local Similarity 100.0%; Pred. No. 5.7e-285;  
Matches 921; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGTCTCAAAATGCTTATAAGCAGGAAACAGCTGATTTTACTACTAGGAATAGTCTTT 60  
Db 1 ATGAGTCTCAAAATGCTTATAAGCAGGAAACAGCTGATTTTACTACTAGGAATAGTCTTT 60  
Qy 61 TTGAACRAGGTAATCTGCARCTCTTTTCGCTCCCAAGCTCCAGTCTGGGACAGT 120  
Db 61 TTGAACRAGGTAATCTGCARCTCTTTTCGCTCCCAAGCTCCAGTCTGGGACAGT 120  
Qy 121 CTGTTTAAAGGTACAGCTTGGAAATATTTTAAACATTTTTCAGTCCGATCTTTGGAGGAGC 180  
Db 121 CTGTTTAAAGGTACAGCTTGGAAATATTTTAAACATTTTTCAGTCCGATCTTTGGAGGAGC 180  
Qy 181 CAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAACAAAGGACAGCATATT 240  
Db 181 CAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAACAAAGGACAGCATATT 240  
Qy 241 TGTGGAGGAGCATCGTCTCACCACAGTGGGTGATCATCGCGGCTCACTGCATGCAAAAC 300  
Db 241 TGTGGAGGAGCATCGTCTCACCACAGTGGGTGATCATCGCGGCTCACTGCATGCAAAAC 300  
Qy 301 AGAAACATTTGTTACTTTTGAATGTTTACTGCTGGAGAGTATGACTTAAAGCAGACAGAC 360  
Db 301 AGAAACATTTGTTACTTTTGAATGTTTACTGCTGGAGAGTATGACTTAAAGCAGACAGAC 360  
Qy 361 CCAGGAGAGCAAACTCTCATTGAACTGTCTATCATATCATCATCATCATCATCATCATCAT 420  
Db 361 CCAGGAGAGCAAACTCTCATTGAACTGTCTATCATATCATCATCATCATCATCATCATCAT 420  
Qy 421 AAACCAATGGACTATGATATTTGCGCTTTTGAAGATGGCTTCCCAATTTGGCCAC 480  
Db 421 AAACCAATGGACTATGATATTTGCGCTTTTGAAGATGGCTTCCCAATTTGGCCAC 480  
Qy 481 TTTGTGGGGCCCATATCTCTCCAGAGCTGCGGGAGCAATTTGAGGCTGTTTATTGTT 540

Db 481 TTTGTGGGGCCCATATCTCTCCAGAGCTGCGGGAGCAATTTGAGGCTGGTTATTGTT 540  
Qy 541 ACAACTGCAGGCTGGGCGCTTAACCTCAAGGTGGCTCTCTCACAGTCTTTCAGGAA 600  
Db 541 ACAACTGCAGGCTGGGCGCTTAACCTCAAGGTGGCTCTCTCACAGTCTTTCAGGAA 600  
Qy 601 GTGAATCTGCCATTTTGGACCTGGGAAGAGTGTGTGGCAGCTCTGTTAACACTAAAGAGG 660  
Db 601 GTGAATCTGCCATTTTGGACCTGGGAAGAGTGTGTGGCAGCTCTGTTAACACTAAAGAGG 660  
Qy 661 CCCATCAGTGGGAAGACCTTTCTTTTGACAGAGTGTGTGGGAGGAGGAGAGCATGTT 720  
Db 661 CCCATCAGTGGGAAGACCTTTCTTTTGACAGAGTGTGTGGGAGGAGGAGAGCATGTT 720  
Qy 721 CAGGAGATTCAGGAGTTCATCTCATGTGCCGGAATAGAAAGGGGCTTGACACTGGCT 780  
Db 721 CAGGAGATTCAGGAGTTCATCTCATGTGCCGGAATAGAAAGGGGCTTGACACTGGCT 780  
Qy 781 GGTGTGACTTCTCTGGGGTTTGGGCTGTGTGGTGGAGGCTGGAGAAACAATGTGAGGAAAGT 840  
Db 781 GGTGTGACTTCTCTGGGGTTTGGGCTGTGTGGTGGAGGCTGGAGAAACAATGTGAGGAAAGT 840  
Qy 841 GATCAAGATCCCTGGGATCTTCACAGACATTTAGTAAAGTCTTCTTGATCCAGCA 900  
Db 841 GATCAAGATCCCTGGGATCTTCACAGACATTTAGTAAAGTCTTCTTGATCCAGCA 900  
Qy 901 CACATCCAAACTGGTAACTAA 921  
Db 901 CACATCCAAACTGGTAACTAA 921

RESULT 2  
ABK12903  
ID ABK12903 standard; cDNA; 1262 BP.  
XX  
AC ABK12903;  
XX  
DT 09-APR-2002 (first entry)  
XX  
XX Human protease PRTS-20 cDNA sequence.  
XX  
XX Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;  
KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;  
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;  
KW cell proliferative disorder; developmental disorder; epilepsy;  
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;  
KW reproductive disorder; endometriosis; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 165..1085  
FT /\*tag= a  
FT /\*product= "Human protease PRTS-20"  
XX  
XX WO200198468-A2.  
XX  
XX 27-DEC-2001.  
XX  
XX 13-JUN-2001; 2001WO-US19178.  
XX  
XX 16-JUN-2000; 2000US-212336P.  
XX 22-JUN-2000; 2000US-213955P.  
XX 29-JUN-2000; 2000US-215396P.  
XX 07-JUL-2000; 2000US-216821P.  
XX 14-JUL-2000; 2000US-218946P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Yue H, Elliott VS, Gandhi AR, Lal P, Au-young J, Tribouley CM;  
PI Deleageane AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;  
PI Walia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;

PI Azimzal Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;  
PI Kallick DA;  
DR WPI: 2002-090437/12.  
DR P-PSDB: AAU74760.  
XX  
XX Twenty one human proteases (referred to as PRS-1 to PRS-21), useful  
PT in the diagnosis, treatment and prevention of gastrointestinal (e.g.  
PT gastritis), cardiovascular (e.g. atherosclerosis) and cell  
PT proliferative (e.g. cancer) disorders -  
XX  
XX Claim 5; Page 175; 177pp; English.  
XX  
XX The present invention relates to twenty one new human proteases,  
CC referred to as PRS-1 to PRS-21. The PRS polynucleotides and  
CC polypeptides of the invention are useful in the diagnosis, treatment and  
CC prevention of gastrointestinal e.g. gastritis, esophageal carcinoma and  
CC Crohn's disease, cardiovascular e.g. atherosclerosis, hypertension and  
CC myocardial infarction, autoimmune/inflammatory e.g. acquired  
CC immunodeficiency syndrome (AIDS), allergies and osteoarthritis, cell  
CC proliferative e.g. cancer, developmental e.g. Duchenne and Becker  
CC muscular dystrophy, epithelial e.g. dermatitis, neurological e.g.  
CC epilepsy and Alzheimer's disease and reproductive e.g. infertility and  
CC endometriosis disorders. Numerous other examples of each disorder are  
CC given in the specification. The present nucleic acid sequence encodes  
CC the human protease PRS-20 protein of the invention.  
XX  
XX Sequence 1262 BP; 354 A; 265 C; 306 G; 337 T; 0 other;  
SQ

Query Match 99.9%; Score 920.2; DB 24; Length 1262;  
Best Local Similarity 99.8%; Pred. No. 6.7e-285;  
Matches 919; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGTCTCAAAATGCTTATAGCAGGAAACAGCTGATTTTACTACTAGGATAGTCTTT 60  
Db 165 ATGAGTCTCAAAATGCTTATAGCAGGAAACAGCTGATTTTACTACTAGGATAGTCTTT 224  
Qy 61 TTTGAACRAGGTAAATCTGCAACTCTTTTCGCTCCCAAGCTCCAGTTGGGCGAGGT 120  
Db 225 TTTGAACRAGGTAAATCTGCAACTCTTTTCGCTCCCAAGCTCCAGTTGGGCGAGGT 284  
Qy 121 CTGGTTAAGGTACAGCTTGGAAATATTTTAAACATTTTCAGTCGCAATTTGGAGGAAGC 180  
Db 285 CTGGTTAAGGTACAGCTTGGAAATATTTTAAACATTTTCAGTCGCAATTTGGAGGAAGC 344  
Qy 181 CAAGTGGAGAGGTTCTATCCCTGGCAGGTATCTCTGAACAAAGCAGCAAGCATATT 240  
Db 345 CAAGTGGAGAGGTTCTATCCCTGGCAGGTATCTCTGAACAAAGCAGCAAGCATATT 404  
Qy 241 TGTGAGGAGAGCATGCTCTACCCAGTGGGTGATCAGCGGGCTCACTGCAATTCGAAC 300  
Db 405 TGTGAGGAGAGCATGCTCTACCCAGTGGGTGATCAGCGGGCTCACTGCAATTCGAAC 464  
Qy 301 AGAACATTTGCTTACTTTTGAATGTTACTGCTGGAGAGTATGACTTAAGCCAGACAGAC 360  
Db 465 AGAACATTTGCTTACTTTTGAATGTTACTGCTGGAGAGTATGACTTAAGCCAGACAGAC 524  
Qy 361 CCAGGAGAGCAACTCTCATTATGAACATGCTCATCATCATCCATCTTCCACCAG 420  
Db 525 CCAGGAGAGCAACTCTCATTATGAACATGCTCATCATCATCCATCTTCCACCAG 584  
Qy 421 AAACCAATGGACTATGATATGCCCTTTTGAAGATGGCTGGAGCCCTCCAAATTTGGCCAC 480  
Db 585 AAACCAATGGACTATGATATGCCCTTTTGAAGATGGCTGGAGCCCTCCAAATTTGGCCAC 644  
Qy 481 TTTGTGGGGCCCATATGCTTCCAGAGCTGCGGGAGCAATTTAGCGCTGGTTTATTTGT 540  
Db 645 TTTGTGGGGCCCATATGCTTCCAGAGCTGCGGGAGCAATTTAGCGCTGGTTTATTTGT 704  
Qy 541 ACAACTGCAGCTGGGGCCGCTTAACCTAAGTGGCGTCTCTCAACAGCTTTCAGGAA 600  
Db 705 ACAACTGCAGCTGGGGCCGCTTAACCTAAGTGGCGTCTCTCAACAGCTTTCAGGAA 764

Qy 601 GTGAATCTGCTTATTTTACCTGGGAGAGTGTGTGGCAGCTCTGTTAACTAAAGAGG 660  
Db 765 GTGAATCTGCTTATTTTACCTGGGAGAGTGTGTGGCAGCTCTGTTAACTAAAGAGG 824  
Qy 661 CCCATCACTGGGAGAGCTTCTTTTGCACAGGTTTCTGATGGAGGAGAGCAGCATGT 720  
Db 825 CCCATCACTGGGAGAGCTTCTTTTGCACAGGTTTCTGATGGAGGAGAGCAGCATGT 884  
Qy 721 CAGGAGATTCAGGAGGTTCACTCATGTGCGGGAATAAGAAAGGCGCTGGACTTGGCT 780  
Db 885 CAGGAGATTCAGGAGGTTCACTCATGTGCGGGAATAAGAAAGGCGCTGGACTTGGCT 944  
Qy 781 GGTGTGACTTCTGGGTTTGGGCTGTGTGAGCTGGAGAAACAATGTGAGGAAAAGT 840  
Db 945 GGTGTGACTTCTGGGTTTGGGCTGTGTGAGCTGGAGAAACAATGTGAGGAAAAGT 1004  
Qy 841 GATCAAGGATCCCTGGGATCTTCACAGACATTTAGTAAGTGTCTTCTGGATCCACGAA 900  
Db 1005 GATCAAGGATCCCTGGGATCTTCACAGACATTTAGTAAGTGTCTTCTGGATCCACGAA 1064  
Qy 901 CACATCCAAACTGGTAATAA 921  
Db 1065 CACATCCAAACTGGTAATAA 1085  
RESULT 3  
AAD09331  
ID AAD09331 standard; cDNA; 1568 BP.  
XX  
AC AAD09331;  
XX  
DT 10-SEP-2001 (first entry)  
XX  
DE Human protease homologue cDNA #4.  
XX  
KW Human; protease homologue; novel human protein; NHP; therapy;  
KW pharmacogenomic application; physiological disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200146407-A1.  
XX  
PD 28-JUN-2001.  
XX  
PF 12-DEC-2000; 2000WO-US33738.  
XX  
PR 23-DEC-1999; 99US-0171566.  
XX  
PA (LEXI-) LEXICON GENETICS INC.  
XX  
PL Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;  
XX WPI: 2001-408641/43.  
XX  
XX Polynucleotide encoding novel human protease homologs, useful for  
PT identifying agonist, antagonist or modifiers or for producing  
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic  
PT applications -  
XX  
PS Disclosure: Page 31; 32pp; English.  
XX  
CC The present sequence is a cDNA encoding novel human protein (NHP),  
CC known as human protease homologue. NHP shares structural similarity  
CC with animal proteases, particularly trypsin-like protease such  
CC as oviductin, plasminogen activator and human plasma kallikrein  
CC precursor. NHP and its cDNA are useful in therapeutic, diagnostic and  
CC pharmacogenomic applications. NHP sequences are useful for identifying  
CC agonists, antagonists and modulators and also for producing antibodies  
CC useful in diagnosis, drug screening, clinical trial monitoring and in  
CC treatment of physiological disorders.  
XX  
SQ Sequence 1568 BP; 452 A; 338 C; 353 G; 423 T; 2 other;

Query Match		99.98;	Score 920.2;	DB 22;	Length 1568;
Best Local Similarity		100.0%;	Pred. No. 7.5e-285;		
Matches 921;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	ATGAGTCTCAAAATGCTTATAAGCAGGAACAAGCTGATTTTACTACTAGGAATAGTCTTT	60		
Db	245	ATGAGTCTCAAAATGCTTATAAGCAGGNACAAGCTGATTTTACTACTAGGAATAGTCTTT	304		
Qy	61	TTTGAACRAGGTAATCTGCAACCTCTTTTCGCTCCCAAGCTCCCAAGTGTGGGACAGAGT	120		
Db	305	TTTGAACRAGGTAATCTGCAACCTCTTTTCGCTCCCAAGCTCCCAAGTGTGGGACAGAGT	364		
Qy	121	CTGGTTAAGGTACAGCCTTGGAAATATTTTAAACATTTTTCAGTCGCATTTCTGGAGGAAGC	180		
Db	365	CTGGTTAAGGTACAGCCTTGGAAATATTTTAAACATTTTTCAGTCGCATTTCTGGAGGAAGC	424		
Qy	181	CAAGTGGAGAAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAAGGCGAAGCATATT	240		
Db	425	CAAGTGGAGAAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAAGGCGAAGCATATT	484		
Qy	241	TGTGGAGGAAGCATCTGCTCACACAGTGGGTGATCACGGCGGCTCACTGCATTGCAAAAC	300		
Db	485	TGTGGAGGAAGCATCTGCTCACACAGTGGGTGATCACGGCGGCTCACTGCATTGCAAAAC	544		
Qy	301	AGAAACATTGTCTACTTTGAATGTTACTGCTGGAGAGTATGACTTAAAGCCAGACAGAC	360		
Db	545	AGAAACATTGTGTCTACTTTGAATGTTACTGCTGGAGAGTATGACTTAAAGCCAGACAGAC	604		
Qy	361	CCAGGAGGCAAACTCTCACTATTGAACTGTCATCATACATCCACATTTCTCCACCAAG	420		
Db	605	CCAGGAGGCAAACTCTCACTATTGAACTGTCATCATACATCCACATTTCTCCACCAAG	664		
Qy	421	AAACCAATGGACTATGATATTGCCCTTTTGAAGATGGCTGGAGCCCTTCCAAATTTGGCCAC	480		
Db	665	AAACCAATGGACTATGATATTGCCCTTTTGAAGATGGCTGGAGCCCTTCCAAATTTGGCCAC	724		
Qy	481	TTTGTGGGGCCCATATGTCTTCCAGAGCTGCGGGAGCAATTTGAGGCTGGTTTATTGT	540		
Db	725	TTTGTGGGGCCCATATGTCTTCCAGAGCTGCGGGAGCAATTTGAGGCTGGTTTATTGT	784		
Qy	541	ACAACCTGCAGGCTGGGGCCGCTTAACTGAAGGTGGCGTCTCTCAACAAGTCTTGCAAGAA	600		
Db	785	ACAACCTGCAGGCTGGGGCCGCTTAACTGAAGGTGGCGTCTCTCAACAAGTCTTGCAAGAA	844		
Qy	601	GTGAATCTGCCATTTTTCAGCTGGGAAGAGTGTGTGGCAGCTCTGTTAACACTATAAGAGG	660		
Db	845	GTGAATCTGCCATTTTTCAGCTGGGAAGAGTGTGTGGCAGCTCTGTTAACACTATAAGAGG	904		
Qy	661	CCCATCAGTGGGAAGACCTTTCTTTGACACAGTTTTCCTGATGGAGGGAGAGACGCATGT	720		
Db	905	CCCATCAGTGGGAAGACCTTTCTTTGACACAGTTTTCCTGATGGAGGGAGAGACGCATGT	964		
Qy	721	CAGGGAGATTCAGGAGGTTCACTCATGTGCCGGAATTAAGAAAGGGCCCTGGACTCTGGCT	780		
Db	965	CAGGGAGATTCAGGAGGTTCACTCATGTGCCGGAATTAAGAAAGGGCCCTGGACTCTGGCT	1024		
Qy	781	GGTGTGACTTCTGGGGTTTGGGCTGTGGTGCAGGCTGGAGAAACAATGTGAGGAAAAGT	840		
Db	1025	GGTGTGACTTCTGGGGTTTGGGCTGTGGTGCAGGCTGGAGAAACAATGTGAGGAAAAGT	1084		
Qy	841	GATCAAGGATCCCTGGGATCTTCACAGACATTTAGTAAAGTGTCTTCTCGATCCACGAA	900		
Db	1085	GATCAAGGATCCCTGGGATCTTCACAGACATTTAGTAAAGTGTCTTCTCGATCCACGAA	1144		
Qy	901	CACATCCAAACTGGTAACATA	921		
Db	1145	CACATCCAAACTGGTAACATA	1165		

RESULT 4  
AAD09329  
ID AAD09329 standard; cDNA; 909 BP.  
XX

AC	AAD09329;	
XX		
DT	10-SEP-2001 (first entry)	
XX		
DE	Human protease homologue cDNA #2.	
XX		
KW	Human; protease homologue; novel human protein; NHP; therapy;	
KW	pharmacogenomic application; physiological disorder; ss.	
XX		
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	1..909
FT		/tag= a
FT		/product= "Human protease homologue #2"
FT	unsure	55..57
FT		/tag= b
FT		/note= "Encodes Arg"
FT	misc_feature	56
FT		/tag= c
FT		/note= "This degenerate base represents a polymorphic site"
FT	unsure	70..72
FT		/tag= d
FT		/note= "Encodes Ala"
FT	misc_feature	70
FT		/tag= e
FT		/note= "This degenerate base represents a polymorphic site"
XX		
PN	WO200146407-A1.	
XX		
PD	28-JUN-2001.	
XX		
PF	12-DEC-2000; 2000WO-US33738.	
XX		
PR	23-DEC-1999; 99US-0171566.	
XX		
PA	(LEXI-) LEXICON GENETICS INC.	
XX		
PI	Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;	
XX		
DR	WPI; 2001-408641/43.	
DR	P-PSDB; AAE04734.	
XX		
PT	Polynucleotide encoding novel human protease homologs, useful for	
PT	identifying agonist, antagonist or modifiers or for producing	
PT	antibodies useful in therapeutic, diagnostic and pharmacogenomic	
PT	applications -	
XX		
PS	Disclosure; Page 29; 32pp; English.	
XX		
CC	The present sequence is a cDNA encoding novel human protein (NHP),	
CC	known as human protease homologue. NHP shares structural similarity	
CC	with animal proteases, particularly trypsin-like protease such	
CC	as oviductin, plasminogen activator and human plasma kallikrein	
CC	precursor. NHP and its cDNA are useful in therapeutic, diagnostic and	
CC	pharmacogenomic applications. NHP sequences are useful for identifying	
CC	agonists, antagonists and modulators and also for producing antibodies	
CC	useful in diagnosis, drug screening, clinical trial monitoring and in	
CC	treatment of physiological disorders.	
XX		
SQ	Sequence 909 BP; 237 A; 193 C; 238 G; 239 T; 2 other;	

Query Match 98.6%; Score 908.2; DB 22; Length 909;  
Best Local Similarity 100.0%; Pred. No. 4e-281;  
Matches 909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	13	ATGCTTATAAGCAGGAAACAGCTGATTTTACTACTAGGAATAGTCTTTTGAACRAGGT	72
DB	1	ATGCTTATAAGCAGGAAACAGCTGATTTTACTACTAGGAATAGTCTTTTGAACRAGGT	60
QY	73	AAATCTGCARCTCTTTCGGTCCCAAGCTCCCAAGTTCGGGCGAGAGTCTGGTTAAGGTA	132

|||||  
Db 61 AAATGTCARCTTTTCGTCGCCAAAGCTCCAGTTGTGGCAGAGTCTGGTAAAGTA 120  
QY 133 CAGCCTTGGAATATTTTAAACATTTTCAGTCGCATTTCTGGAGGAGCCAAAGTGGAAG 192  
Db 121 CAGCCTTGGAATATTTTAAACATTTTCAGTCGCATTTCTGGAGGAGCCAAAGTGGAAG 180  
QY 193 GGTTCCTATCCCTGGCAGGTATCTCTGAAACAAAGCAGAGCATATTTGTGGAGGAGC 252  
Db 181 GGTTCCTATCCCTGGCAGGTATCTCTGAAACAAAGCAGAGCATATTTGTGGAGGAGC 240  
QY 253 ATCGCTCACACAGTGGGTGATCAGCGCGGCTCACTGCTGATTCGAAACAGAAACATTTG 312  
Db 241 ATCGCTCACACAGTGGGTGATCAGCGCGGCTCACTGCTGATTCGAAACAGAAACATTTG 300  
QY 313 TCTACTTTTCAATGTTACTGCTGGAGAGTATGACTTAAGCCAGACAGCCAGGAGAGCA 372  
Db 301 TCTACTTTTCAATGTTACTGCTGGAGAGTATGACTTAAGCCAGACAGCCAGGAGAGCA 360  
QY 373 ACTCTACTATTGAAACTGCTCATCATCATCATCATCATCTTCCACCAAGAAACCAATGGAC 432  
Db 361 ACTCTACTATTGAAACTGCTCATCATCATCATCATCTTCCACCAAGAAACCAATGGAC 420  
QY 433 TATGATATTGCCCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCACTTTGTGGGGCC 492  
Db 421 TATGATATTGCCCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCACTTTGTGGGGCC 480  
QY 493 ATATGCTTCCAGAGCTGCGGGAGCAATTTGAGGCTGGTTTATTGTACAACCTGCAGGC 552  
Db 481 ATATGCTTCCAGAGCTGCGGGAGCAATTTGAGGCTGGTTTATTGTACAACCTGCAGGC 540  
QY 553 TGGGGCCGCTTAAGTGAAGGTGGCGTCTCTCACAAAGTCTTGCAAGAGTGAATCTGCCT 612  
Db 541 TGGGGCCGCTTAAGTGAAGGTGGCGTCTCTCACAAAGTCTTGCAAGAGTGAATCTGCCT 600  
QY 613 ATTTGACCTGGAGAGTGTGGCAGCTCTGTTAAACACTAAAGAGGCCCATCAGTGG 672  
Db 601 ATTTGACCTGGAGAGTGTGGCAGCTCTGTTAAACACTAAAGAGGCCCATCAGTGG 660  
QY 673 AAGACCTTCTTTGACAGGTTTCTGATGAGGAGAGAGCGATGTTCAGGAGAGATTCA 732  
Db 661 AAGACCTTCTTTGACAGGTTTCTGATGAGGAGAGAGCGATGTTCAGGAGAGATTCA 720  
QY 733 GGAGGTTCATCATGTGCCGGAATAAGAAAGGGGCGCTGCACTCTGCTGGTGTGACTTCC 792  
Db 721 GGAGGTTCATCATGTGCCGGAATAAGAAAGGGGCGCTGCACTCTGCTGGTGTGACTTCC 780  
QY 793 TGGGGTTTGGGCTGTGGTGGAGCTGGAGAAACAAATGTCAGGAAAGTGTCAAGGATCC 852  
Db 781 TGGGGTTTGGGCTGTGGTGGAGCTGGAGAAACAAATGTCAGGAAAGTGTCAAGGATCC 840  
QY 853 CCTGGGATCTTCACAGACATTAAGTAAAGTCTTCTGATCCACCAACACATCCAAACT 912  
Db 841 CCTGGGATCTTCACAGACATTAAGTAAAGTCTTCTGATCCACCAACACATCCAAACT 900  
QY 913 GGTAACTAA 921  
Db 901 GGTAACTAA 909  
  
RESULT 5  
ID ABK31790  
AC ABK31790 standard; DNA; 1671 BP.  
XX  
AC ABK31790;  
XX  
DT 23-APR-2002 (first entry)  
XX  
DE DNA encoding novel human protease #47.  
XX  
KW Human; protease; cancer; immune-related disorder; cardiovascular disease;  
KW neuronal-associated disease; metabolic disorder; inflammatory disorder;  
KW nervous system disorder; sexual dysfunction; pain; mood disorder;

KW hypertension; psychotic disorder; neurological disorder; dyskinesia;  
KW viral infection; human immunodeficiency virus; HIV; non-viral infection;  
XX ocular disease; cytostatic; gene; ds.  
OS Homo sapiens.  
XX WO200200860-A2.  
PN 03-JAN-2002.  
XX  
PF 26-JUN-2001; 2001WO-US20171.  
XX  
PR 26-JUN-2000; 2000US-214047P.  
PI (SUC-) SUGEN INC.  
XX  
PI Plowman G, Whyte D, Sudarsanam S, Manning G, Caenepeel S;  
PI Charydzak G;  
XX  
XX WPI; 2002-139913/18.  
DR P-PSDB; AAU82748.  
XX  
XX Nucleic acids encoding novel human proteases, useful for useful for  
PT treating diseases and disorders such as cancers, immune-related  
PT diseases and disorders, cardiovascular disease (e.g. restenosis) and  
PT inflammatory disorders  
XX  
PS Claim 26; Fig 100-PP; 313pp; English.  
XX  
CC The present invention relates to the isolation of novel human  
CC proteases, and the nucleic acids encoding them. The sequences of  
CC the invention are useful for treating diseases and disorders such as  
CC cancers (e.g. breast, colon, lung), immune-related diseases and disorders  
CC (e.g. inflammatory diseases and asthma), cardiovascular diseases  
CC (e.g. restenosis and coronary thrombosis), brain or neuronal-associated  
CC diseases, metabolic disorders (e.g. diabetes, obesity), inflammatory  
CC disorders (e.g. rheumatoid arthritis and psoriasis), central or  
CC peripheral nervous system diseases, migraines, pain, sexual dysfunction,  
CC mood disorders, attention disorders, cognition disorders, hypotension,  
CC hypertension, psychotic disorders, neurological disorders  
CC (e.g. Alzheimer's disease, Parkinson's disease) and dyskinesias.  
CC The nucleic acids and polypeptides are also useful for treating viral  
CC infections caused by human immunodeficiency virus (HIV), and non-viral  
CC infections such as ocular disease (e.g. glaucoma) and macular  
CC degeneration. ABK31744-ABK31802 represent DNA sequences encoding for  
CC the novel human proteases of the invention.  
XX  
SQ Sequence 1671 BP; 451 A; 369 C; 404 G; 447 T; 0 other;

Query Match 83.7%; Score 771.2; DB 24; Length 1671;  
Best Local Similarity 99.6%; Pred. No. 6e-237;  
Matches 781; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 ATGAGTCTCAAAATGCTTATAGCAGGACACAGCTGATTTTACTAGTAATAGTCTTT 60  
Db 1 ATGAGTCTCAAAATGCTTATAGCAGGACACAGCTGATTTTACTAGTAATAGTCTTT 60  
QY 61 TTTGAACRAGGTAAATCTGCARCTCTTCGCTCCCAAGCTCCAGTTGTGGGAGAGT 120  
Db 61 TTTGAACRAGGTAAATCTGCARCTCTTCGCTCCCAAGCTCCAGTTGTGGGAGAGT 120  
QY 121 CTGGTTAAGGTACAGCTTTGGAATTTATTTAAACATTTTTCAGTCGCATTTCTTGGAGGAAGC 180  
Db 121 CTGGTTAAGGTACAGCTTTGGAATTTATTTAAACATTTTTCAGTCGCATTTCTTGGAGGAAGC 180  
QY 181 CAAGTGGAGAGGGTTCCCTATCCCTGGCAGGTATCTCTGAAACAAAGGAGGAGCATATT 240  
Db 181 CAAGTGGAGAGGGTTCCCTATCCCTGGCAGGTATCTCTGAAACAAAGGAGGAGCATATT 240  
QY 241 TGTGGAGGAGGATCTCTCTCACCACAGTGGGTGATCAGCGGCTCCTCCTGCTTGGCAAC 300  
Db 241 TGTGGAGGAGGATCTCTCTCACCACAGTGGGTGATCAGCGGCTCCTCCTGCTTGGCAAC 300

QY 301 AGAACATTTGTCTACTTTGAATGTTACTGCTGGAGAGTATGACTTTAAGCCAGACAGAC 360  
DB 301 AGAACATTTGTCTACTTTGAATGTTACTGCTGGAGAGTATGACTTTAAGCCAGACAGAC 360  
QY 361 CCAGGAGCAAACTCTACATATTGAACCTGTCTATCATATCATATCCACATTTCTCCACCAAG 420  
DB 361 CCAGGAGCAAACTCTACATATTGAACCTGTCTATCATATCATATCCACATTTCTCCACCAAG 420  
QY 421 AAACCAATGGACTATGATATGCTTTCCTTTTGAAGATGGCTGGAGCCTTCCCAATTTGGCCAC 480  
DB 421 AAACCAATGGACTATGATATGCTTTCCTTTTGAAGATGGCTGGAGCCTTCCCAATTTGGCCAC 480  
QY 481 TTTGTGGGGCCCATATGCTTCTCCAGAGCTGCGGAGCAATTTGAGGCTGGTTTATTATTGT 540  
DB 481 TTTGTGGGGCCCATATGCTTCTCCAGAGCTGCGGAGCAATTTGAGGCTGGTTTATTATTGT 540  
QY 541 ACAACTGAGAGCTGGGCGCGCTTAACTGAAGGTGGCGTCTCTCTCACAAAGTCTTTCAGAGAA 600  
DB 541 ACAACTGAGAGCTGGGCGCGCTTAACTGAAGGTGGCGTCTCTCTCACAAAGTCTTTCAGAGAA 600  
QY 601 GTGAATCTGCTATTTTACACTGGGAGAGCTGTGTGGCAGCTCTGTTAAACACTTAAAGAGG 660  
DB 601 GTGAATCTGCTATTTTACACTGGGAGAGCTGTGTGGCAGCTCTGTTAAACACTTAAAGAGG 660  
QY 661 CCCATCAGTGGGAAGACCTTTCTTTTGCACAGGTTTTCCTGATGGAGGAGAGAGCGCATGT 720  
DB 661 CCCATCAGTGGGAAGACCTTTCTTTTGCACAGGTTTTCCTGATGGAGGAGAGAGCGCATGT 720  
QY 721 CAGGAGATTCAGGAGGTTCACTCATGTGCGGGAATTAAGAAAGGGGCTTGGACTCTGGC 779  
DB 721 CAGGAGATTCAGGAGGTTCACTCATGTGCGGGAATTAAGAAAGGGGCTTGGACTCTGGC 780  
QY 780 TGGT 783  
DB 781 TGGT 784  
RESULT 6  
AAH98934  
ID AAH98934 standard; cDNA; 637 BP.  
XX AC AAH98934;  
XX DT 12-OCT-2001 (first entry)  
XX Toad EST-derived coding sequence SEQ ID NO: 791.  
XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
KW diagnostics; forensic test; gene mapping; genetic disorder;  
KW biodiversity; gene therapy; nutrition; ss.  
OS Xenopus laevis.  
XX WO200154477-A2.  
XX PD 02-AUG-2001.  
XX PF 25-JAN-2001; 2001WO-US02687.  
XX PR 25-JAN-2000; 2000US-0491404.  
XX PR 17-JUL-2000; 2000US-0617746.  
XX PR 03-AUG-2000; 2000US-0631451.  
XX PR 15-SEP-2000; 2000US-0663870.  
XX PA (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
PI Cao Y, Drmanac RA, Zhang J, Werhman T;  
XX WPI; 2001-476164/51.  
DR P-PSDB; AAM24275.  
XX

PT Isolated polypeptide for treatment of diseases, diagnostics, raising  
PT antibodies and research use -  
XX Claim 1; Page 673-674; 1275pp; English.  
PS The present invention provides the protein and coding sequences of novel  
XX proteins from a variety of organisms, including human, dog, cat, horse,  
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)  
CC from the organism of interest. They can be used in diagnostics,  
CC forensics, gene mapping, identification of mutations, to assess  
CC biodiversity and for nutritional purposes. The present sequence is a cDNA  
XX of the invention.  
SQ Sequence 637 BP; 172 A; 133 C; 141 G; 191 T; 0 other;

Query Match 59.3%; Score 546.2; DB 22; Length 637;  
Best Local Similarity 99.6%; Pred. No. 7.3e-165;  
Matches 545; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAGTCTCAAAATGCTTATAAGCAGCAAGCTGATTTTACTAGGAAATAGTCTTT 60  
DB 91 ATGAGTCTCAAAATGCTTATAAGCAGCAAGCTGATTTTACTAGGAAATAGTCTTT 150  
QY 61 TTTGAACRAGTAATCTGCARCTCTTTCGCTCCCAAGCTCCAGTTGTGGCAGAGT 120  
DB 151 TTTGAACGAGTAATCTGCACATCTTTCGCTCCCAAGCTCCAGTTGTGGCAGAGT 210  
QY 121 CTGTTTAAAGTACAGCCTTGGAAATTTTAAACATTTTCACTCGCATTTTGGAGGAAGC 180  
DB 211 CTGTTTAAAGTACAGCCTTGGAAATTTTAAACATTTTCACTCGCATTTTGGAGGAAGC 270  
QY 181 CAAGTGAGAGAGGTTCCATCCCTGCAGGATATCTCTGAACAAAGCAGACATATT 240  
DB 271 CAAGTGAGAGAGGTTCCATCCCTGCAGGATATCTCTGAACAAAGCAGACATATT 330  
QY 241 TGTGGAGGAAGCATCGTCTCACCACAGTGGGTGATCAGCGGGCTCACTGCATTGCAAAAC 300  
DB 331 TGTGGAGGAAGCATCGTCTCACCACAGTGGGTGATCAGCGGGCTCACTGCATTGCAAAAC 390  
QY 301 AGAACATTTGTCTACTTTTGAATGTTTACTGCTGGAGAGTATGACTTTAAGCCAGACAGAC 360  
DB 391 AGAACATTTGTCTACTTTTGAATGTTTACTGCTGGAGAGTATGACTTTAAGCCAGACAGAC 450  
QY 361 CCAGGAGCAAACTCTACATATTGAACCTGTCTATCATATCCACATTTCTCCACCAAG 420  
DB 451 CCAGGAGCAAACTCTACATATTGAACCTGTCTATCATATCCACATTTCTCCACCAAG 510  
QY 421 AAACCAATGGACTATGATATGCTTTCCTTTTGAAGATGGCTGGAGCCTTCCCAATTTGGCCAC 480  
DB 511 AAACCAATGGACTATGATATGCTTTCCTTTTGAAGATGGCTGGAGCCTTCCCAATTTGGCCAC 570  
QY 481 TTTGTGGGGCCCATATGCTTCTCCAGAGCTGCGGAGCAATTTGAGGCTGGTTTATTATTGT 540  
DB 571 TTTGTGGGGCCCATATGCTTCTCCAGAGCTGCGGAGCAATTTGAGGCTGGTTTATTATTGT 630  
QY 541 ACAACTG 547  
DB 631 ACAACTG 637

RESULT 7  
AAD09330  
ID AAD09330 standard; cDNA; 495 BP.  
XX AC AAD09330;  
XX DT 10-SEP-2001 (first entry)  
XX Human protease homologue cDNA #3.  
XX Human; protease homologue; novel human protein; NHP; therapy;  
KW pharmacogenomic application; physiological disorder; ss.  
KW

XX Homo sapiens.  
OS Key Location/Qualifiers  
FH 1..495  
FT CDS  
FT /\*tag= a  
FT /product= "Human protease homologue #3"  
XX WO200146407-A1.  
XX 28-JUN-2001.  
XX 12-DEC-2000; 2000WO-US33738.  
XX 23-DEC-1999; 99US-0171566.  
XX (LEXI-) LEXICON GENETICS INC.  
XX Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;  
XX WPI; 2001-408641/43.  
XX P-PSDB; AAE04735.  
XX Polynucleotide encoding novel human protease homologs, useful for  
PT identifying agonist, antagonist or modifiers or for producing  
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic  
PT applications -  
XX  
XX Disclosure; Page 30; 32pp; English.  
XX The present sequence is a cDNA encoding novel human protein (NHP),  
CC known as human protease homologue. NHP shares structural similarity  
CC with animal proteases, particularly trypsin-like protease such  
CC as oviductin, plasminogen activator and human plasma kallikrein  
CC precursor. NHP and its cDNA are useful in therapeutic, diagnostic and  
CC pharmacogenomic applications. NHP sequences are useful for identifying  
CC agonists, antagonists and modulators and also for producing antibodies  
CC useful in diagnosis, drug screening, clinical trial monitoring and in  
CC treatment of physiological disorders.  
XX  
XX Sequence 495 BP; 117 A; 101 C; 147 G; 130 T; 0 other;  
SQ  
Query Match 53.7%; Score 495; DB 22; Length 495;  
Best Local Similarity 100.0%; Pred. No. 1.8e-148;  
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 427 ATGGACTATGATATGCGCTTTTGAAGATGGCTGGAGCCCTTCAATTTGGCCACTTTGTG 486  
DB 1 ATGGACTATGATATGCGCTTTTGAAGATGGCTGGAGCCCTTCAATTTGGCCACTTTGTG 60  
QY 487 GGGCCCATATGCTTCCAGAGCTGGGAGCAATTTGAGGCTGGTTTATTCTACAAC 546  
DB 61 GGGCCCATATGCTTCCAGAGCTGGGAGCAATTTGAGGCTGGTTTATTCTACAAC 120  
QY 547 GCAGGCTGGGGCGGCTTAAGTGGCGTCTCTCAAGCTCTTGCAGGAGTGAAT 606  
DB 121 GCAGGCTGGGGCGGCTTAAGTGGCGTCTCTCAAGCTCTTGCAGGAGTGAAT 180  
QY 607 CTGCTATTTTGACCTGGGAAGAGTGTGGGAGCTCTGTTAACTAAAGAGGCCCATC 666  
DB 181 CTGCTATTTTGACCTGGGAAGAGTGTGGGAGCTCTGTTAACTAAAGAGGCCCATC 240  
QY 667 AGTGGGAAGACCTTTTTCACAGCTTTTCCATGATGGAGGGAGACCGCATCTCAGGA 726  
DB 241 AGTGGGAAGACCTTTTTCACAGCTTTTCCATGATGGAGGGAGACCGCATCTCAGGA 300  
QY 727 GATTGAGGAGTTTCACTCATGTGCCGGAATAAGAAAGGGCCCTGGACTCTGGCTGGTGTG 786  
DB 301 GATTGAGGAGTTTCACTCATGTGCCGGAATAAGAAAGGGCCCTGGACTCTGGCTGGTGTG 360  
QY 787 ACTTCTCGGGGTTTGGGCTGTGCTGAGGCTGGAGAAACAATGTGAGGAAAGTATCAA 846  
DB 361 ACTTCTCGGGGTTTGGGCTGTGCTGAGGCTGGAGAAACAATGTGAGGAAAGTATCAA 420

QY 847 GGATCCCTGGGATCTTACAGACATTAGTAAAGTCTTTCCTGGATCCACGAACATC 906  
DB 421 GGATCCCTGGGATCTTACAGACATTAGTAAAGTCTTTCCTGGATCCACGAACATC 480  
QY 907 CAAACTGGTAACATA 921  
DB 481 CAAACTGGTAACATA 495  
RESULT 8  
AAC05816  
ID AAC05816 standard; cDNA; 324 BP.  
XX AAC05816;  
XX 06-OCT-2000 (first entry)  
XX Human secreted protein 5' EST, SEQ ID NO: 9891.  
DE Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KW gene therapy; chromosome mapping; ss.  
XX Homo sapiens.  
OS  
XX EP1033401-A2.  
XX 06-SEP-2000.  
XX 21-FEB-2000; 2000EP-0200610.  
XX 26-FEB-1999; 99US-0122487.  
XX (GEST ) GENSET.  
XX Dumas Milne Edwards J, Duclert A, Giordano J;  
XX WPI; 2000-500381/45.  
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for  
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for  
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -  
XX  
XX Claim 1; SEQ ID 9891; 71pp + CD-ROM; English.  
XX The present sequence is one of a large number of 5' ESTs derived from  
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively  
CC identified within the present sequence. The 5' ESTs were prepared from  
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST  
CC sequences usually correspond mainly to the 3' untranslated region (UTR)  
CC of the mRNA because they are often obtained from oligo-dT primed cDNA  
CC libraries. Such ESTs are not well suited for isolating cDNA sequences  
CC derived from the 5' ends of mRNAs and even in those cases where longer  
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.  
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be  
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used  
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.  
CC They are used to obtain upstream regulatory sequences and to design  
CC expression and secretion vectors.  
XX  
SQ Sequence 324 BP; 91 A; 61 C; 74 G; 94 T; 4 other;  
Query Match 28.2%; Score 260; DB 21; Length 324;  
Best Local Similarity 95.8%; Pred. No. 4.9e-73;  
Matches 272; Conservative 5; Mismatches 5; Indels 2; Gaps 1;  
QY 1 ATGAGTCTCAAAATGCTTATAAGCAGCAACAGCTGATTTTACTACTAGGAATAGTCTTT 60  
DB 41 ATGAGTCTCAAAATGCTTATAAGCAGCAACAGCTGATTTTACTACTAGGAATAGTCTTT 100  
QY 61 TTTGAACRAGGTAATCTGCARCTTTTCGCTCCCCAAAGCTCCCAAGTT--GTGGGCGA 118  
DB 101 TTTGAACRAGGTAATCTGCARCTTTTCGCTCCCCAAAGCTCCCAAGTTTGTGGCGCARA 160





CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 4729 BP; 1652 A; 1025 C; 960 G; 1092 T; 0 other;

Query Match 17.7%; Score 163.2; DB 23; Length 4729;  
Best Local Similarity 51.7%; Pred. No. 2.6e-41;  
Matches 400; Conservative 0; Mismatches 368; Indels 6; Gaps 1;  
QY 148 TTAAACATTTTCAGTCGCATTTCTTGGAGAGCCCAAGTGGAGAGGGTTCTCTATCCCTGG 207  
DB 2590 TCTAGATTTCTCTAGAAATTAGTAGTTGGAGAAATTCACAGTCACTGGACATCCATGG 2649  
QY 208 CAGGTATCTCTGAACAAAGGAGAGCAATATTTCTGGAGGAAGCATGCTCTCACACAG 267  
DB 2650 CAGGTCTCCCTAAATCAGATGAGCACCACCTTCTGTGGAGGAAGCTGATTCAAGAAGAT 2709  
QY 268 TGGGTGATCAGCGCGCTCAGTGCATTCGACACAGA-----AACATTGTGCTACTTTG 321  
DB 2710 CGGTGTTTACAGCAGCACACTGCGCTGGACGCCCTCAGTGAGAGCAGCTGAAGAATATA 2769  
QY 322 AATGTACTCTGGAGAGTATGACTTAAGCCAGACAGACCCAGGAGAGCAAACTCTCACT 381  
DB 2770 ACTGTGACTTCTGGGAGTACAGCCCTTCTCAGAGGATAACCAAGACAGAAATATCTCT 2829  
QY 382 ATTGAACATGTCATCATACATCCACATTTCTCCACACAGAAACCAATGACTATGATATT 441  
DB 2830 GTCTCAAAATATTATCCCATCTCTGAATACACAGCCGTAATATATGATGCTCTGATATT 2889  
QY 442 GCCCTTTTGAAGATGGCTGGAGCCCTTCCAAATTTGGCCACTTTGTGGGCCCATATGCTTT 501  
DB 2890 GCACGTGCTGTATCTAAAACACAAAGTCAAGTTTGGAAATGCTGTTTCAGCCAACTCTGCTT 2949  
QY 502 CCAGAGCTGCGGGAGCAATTTGAGGCTGCTTTTATTGTTACAACTGCAGGCTGGGGCCGC 561  
DB 2950 CTGTACAGCGATGATTAAGTTGACACAGGAATCTTTGCTATCCAGTGGATGGGGCAAG 3009  
QY 562 TTAACGAAGTGGCGTCTCTCAACAGTCTTTGCAAGGAAGTAATCTCCCTATTTTGGCC 621  
DB 3010 ATTTCACAAACATCAGATATTTCAAAATGCTCTACAAGAAATGGAACCTCCCATCATGAT 3069  
QY 622 TGGGAAGAGTGTGGCAGCTCTGTAACTAAGAGAGCCCATCAGTGGGAGACCTTT 681  
DB 3070 GACAGAGGCTGAATACTGTGCTCAAGAGCATGAACCTCCCTCCCTGGGAGGAGCATG 3129  
QY 682 CTTTCACAGGTTTTCCTGATGGAGGAGAGACGATGTCAGGGAGATTCAGGAGGTTCA 741  
DB 3130 CTGTGCTGGCTTCCCTGATTTGGGATGAGCGCTGCCAGGGGAGCTCTGGAGGACCA 3189  
QY 742 CTCATGTCCCGAATAAGAAAGGGCCCTGGACTCTGGCTGGTGTGACTTCTCTGGGGTTTG 801  
DB 3190 CTGGTTTGTAGAAGAGGTGGTGAATCTGGATTTCTGTGGGATAACTTCTCTGGGTAGCT 3249  
QY 802 GCCTGTGTCGAGGCTGGAGAAACAAATGTGAGGAAAGTATCAAGGATCCCTGGGATC 861  
DB 3250 GGTGTGTGGAGGTTCAAGTTCCCGCTAAGAAACAAACCACTGTGAAGGCATCACTTGGCAT 3309

QY 862 TTACAGACATTAGTAAAGTGCTTTCTGGATCCAGACACATCCAACTGGT 915  
DB 3310 TTCTCCAAAGTCTCTGAGTTGATGATTTATCACTCAAAACCTGTTACAGGT 3363  
RESULT 12  
AAS74405  
ID AAS74405 standard; cDNA; 4729 BP.  
XX AAS74405;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #10209.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US08631.  
XX  
PR 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR WPI; 2001-639362/73.  
DR P-PSDB; ABG10218.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
PS Claim 1; SEQ ID No 10209; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 4729 BP; 1652 A; 1025 C; 960 G; 1092 T; 0 other;

Query Match 17.7%; Score 163.2; DB 23; Length 4729;  
Best Local Similarity 51.7%; Pred. No. 2.6e-41;  
Matches 400; Conservative 0; Mismatches 368; Indels 6; Gaps 1;  
QY 148 TTAAACATTTTCAGTCGCATTTCTTGGAGAGCCCAAGTGGAGAGGGTTCTCTATCCCTGG 207

Db 2590 TCTAGATTCTCTCTAGAAATAGTAGTTGGAGAAATTCACAGTGACTGGACATCCATGG 2649  
Qy 208 CAGGTATCTCTGAACAAGCAGAGCATATTTGTGGAGAACATCGTCTCACCACAG 267  
Db 2650 CAGGTCTCCCTTAAATACAGATGAGCACCACCTCTGTGGAGAACCTTGATTCAAGAAGAT 2709  
Qy 268 TGGGTGATCAGCGCGCTCACTGCAATGCAAAACAGA-----AACATTGTGTCTACTTTG 321  
Db 2710 CGGGTTGTACAGCAGCAGCTGCTGACAGCCCTCAGTGAGAACAGCTGAAGAAATATA 2769  
Qy 322 AATGTACTCTGGAGAGATGACTTAAGCCAGACAGCACCAGGAGCAAACTCTACT 381  
Db 2770 ACTGTGACTCTTGGGAGCTACAGCCCTCTTTCAGAGGATTAAGCAAGACAGAAATATCTCT 2829  
Qy 382 ATTGAACATGTCATCATACATCCACATTTCTCCACCAGAAACCAATGGCATATGATATT 441  
Db 2830 GTCTCAAAAATATTACCCATCTCTGAATACACAGCCGTGAATATAGTCCCTGATATT 2889  
Qy 442 GCCCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCACTTTGTGGGCCCATATGTCTT 501  
Db 2890 GCAGTCTGTATCTAAACACAAAGTCAAGTTTGGAAATGCTGTTCACCAATCTGTCTT 2949  
Qy 502 CCAGAGCTCGGGAGCAATTTGAGGCTGTTTATTTGTACAACATGCAAGGCTGGGGCCGC 561  
Db 2950 CCTGACAGCGATGATAAAGTTGAACCCAGGAATCTCTTGTCTTATCCAGTGGATGGGCAAG 3009  
Qy 562 TTAAGTGAAGTGGGCTCTCTCAAGCTCTGCAAGAGTGAATCTCCCTATTTTGACC 621  
Db 3010 ATTTCCAAAACATCAGAAATATTCAATGTCTTACAGAAATGGAATTTCCCATCATGAT 3069  
Qy 622 TGGGAAGAGTGTGTGGCAGCTCTGTAAACACTAAAGAGGCCCATCAGTGGGAAGACCTTT 681  
Db 3070 GACAGAGCGTGTAACTATGTCTCAAGAGCATGAACCTCCCTCCCTGGGAAGGACCATG 3129  
Qy 682 CTTTCACAGGTTTTCCTGTAGGAGGAGAGACGATGTCAAGGAGATTCAGGAGTTCA 741  
Db 3130 CTGTGTGTGCTTCCCTGATTGGGGAATGGAACGCTCCAGGGGGACTCTGGAGGACCA 3189  
Qy 742 CTCATGTGCGGGAATAAGAAAGGGCGCTGGACTCTGGCTGTGTGACTTCTCTGGGGTTG 801  
Db 3190 CTGGTTGTAGAGAGGTGGTGAATCTGGATCTTGTCTGGGATAACTTCTGGGTAGCT 3249  
Qy 802 GCGTGTGTGCGAGCTGGAGAACAAATGTGAGGAAAAGTATCAAGGATCCCTCGGATC 861  
Db 3250 GCTTGTGTGAGGTTAGTTCCTGTAAGAAACACCACTGTGAAGGCATCACTTGGCAT 3309  
Qy 862 TTCACAGACATAGTAAAGTCTTCTCTGGATCCACGACACATCCAAACTGGT 915  
Db 3310 TTCTCCAAAGTGTCTGAGTTGATGATTATATCACTCAAAACCTGTTACAGGT 3363

RESULT 13

AA578775  
ID AA578775 standard; cDNA: 4729 BP.  
XX  
AC AA578775;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #14579.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US08631.  
XX

PR 31-MAR-2000; 2000US-0540217.  
XX 23-AUG-2000; 2000US-0649167.  
PA (HYSE-) HYSEQ INC.  
XX Drmanac RT, Liu C, Tang YT;  
PI WPI: 2001-639362/73.  
XX P-PSDB; ABG14588.  
DR  
DR  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
XX biodiversity.  
PS Claim 1; SEQ ID No 14579; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: the sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX

Sequence 4729 BP; 1652 A; 1025 C; 960 G; 1092 T; 0 other;

Query Match 17.7%; Score 163.2; DB 23; Length 4729;  
Best Local Similarity 51.7%; Pred. No. 2.6e-41;  
Matches 400; Conservative 0; Mismatches 368; Indels 6; Gaps 1;  
Qy 148 TTTAATTTTTCAGTCGCAATCTTGGAGAACCCAAAGTGGAGAGGTTCTCTATCCCTGG 207  
Db 2590 TCTAGATTCTCTCTAGAAATAGTAGTTGGAGAAATTCACAGTCACTGGACATCCATGG 2649  
Qy 208 CAGGTATCTCTGAACAAGCAGAGCATATTTGTGGAGGAAGCATCGTCTCACCACAG 267  
Db 2650 CAGGTCTCCCTTAAATCAGATGAGCACCACCTCTCTGGAGGAAGCTTGATTCAAGAAGAT 2709  
Qy 268 TGGGTGATCAGCGCGCTCACTGCAATGCAAAACAGA-----AACATTGTGTCTACTTTG 321  
Db 2710 CGGGTTGTACAGCAGCAGCAGCTGCTGACAGCCCTCAGTGAGAACAGCTGAAGAAATATA 2769  
Qy 322 AATGTACTCTGGAGAGTATGACTTAAGCCAGACAGACCCAGGAGCAAACTCTCACT 381  
Db 2770 ACTGTGACTCTTGGGAGTACAGCCCTCTTTCAGAGGATTAAGCAAGACAGAAATATCTCT 2829  
Qy 382 ATTGAACATGTCATCATACATCCACATTTCTCCACCAAGAAACCAATGGCATATGATATT 441  
Db 2830 GTCTCAAAAATATTACCCATCTCTGAATACACAGCCGTGAATATAGTCTCTGATATT 2889  
Qy 442 GCCCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCACTTTGTGGGCCCATATGTCTT 501  
Db 2890 GCAGTCTGTATCTTAAACACAAAGTCAAGTTTGGAAATGCTTTCAGCCCAATGTCTT 2949  
Qy 502 CCAGAGCTCGGGAGCAATTTGAGCGTGTGTTTATTTGTACAACTGCAGGCTGGGGCCGC 561  
Db 2950 CCTGACAGGATGATAAAGTTGAACCCAGGAATCTTGTCTTATCCAGTGGATGGGCAAG 3009



ID AB086158 standard; DNA; 3096 BP.

AC AB086158;

DT 10-SEP-2002 (first entry)

DE Novel human gene. SEQ ID 29.

XX Human; cytostatic; vulnery; antiarteriosclerotic; antiparkinsonian;  
KW neurotropic; neuroprotective; immunosuppressive; haemostatic;  
KW antiinflammatory; cardiant; antiulcer; virucide; antithyroid;  
KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;  
KW wound healing disorders; atherosclerosis; Parkinson's disease;  
KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;  
KW inflammation; neoplastic disease; nervous system disorder;  
KW cardiovascular disorders; pancreatitis; respiratory disorder;  
KW hyperproliferation; systemic autoimmune disease; hyper-immunity;  
KW developmental abnormality; gastrointestinal ulceration; neuropathy;  
KW haematological disease; metabolic disease; sperm dysfunction;  
KW thyroid disorder; hypothyroidism; brain damage; colitis;  
KW cone photo- transduction deficiency; neurological disease; stroke;  
KW angiogenesis; ovulation disorder; spinal cord; thyroid gland; heart;  
KW trachea; thymus; lymph node; muscular system; obesity; anorexia;  
KW growth abnormality; precocious puberty; gene; ss.

XX Homo sapiens.

OS W0200250105-A1.

PN 27-JUN-2002.

PD 17-DEC-2001; 2001WO-US49232.

PF 19-DEC-2000; 2000US-256710P.

PR 20-DEC-2000; 2000US-257048P.

PR 09-JAN-2001; 2001US-260482P.

PR 30-JAN-2001; 2001US-264922P.

PR 06-FEB-2001; 2001US-266797P.

PR 19-MAR-2001; 2001US-276988P.

PR 04-APR-2001; 2001US-281535P.

PR 08-MAY-2001; 2001US-289622P.

XX (SMIK ) SMITHKLINE BEECHAM CORP.

PA (SMIK ) SMITHKLINE BEECHAM PLC.

PA (GLAX ) GLAXO GROUP LTD.

XX Agarwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y;

PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;

XX WPI; 2002-508784/54.

DR P-PSDB; ABP60993.

XX Secreted proteins and polynucleotides useful as vaccines for preventing  
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,  
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune  
PT disorder -

XX Claim 2(a); Page 245-246; 335pp; English.

XX The invention relates to an isolated polypeptide with signal sequences  
CC which allow it to be secreted extracellularly or membrane associated.  
CC The activity of polypeptides of the invention may be described as,  
CC cytostatic, vulnery, antiarteriosclerotic, antiparkinsonian, neurotropic,  
CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,  
CC cardiant, antiulcer, virucide, antithyroid, cerebroprotective, anorectic,  
CC and metabolic. Polypeptides and polynucleotides of the invention are  
CC useful in the treatment, or as a vaccine in the prevention of, cancer,  
CC wound healing disorders, infection, atherosclerosis, parkinson's disease  
CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,  
CC inflammation, neoplastic diseases, nervous system related disorders and  
CC cardiovascular disorders, pancreatitis, respiratory disorder,  
CC hyperproliferation, systemic autoimmune disease, hyper-immunity,  
CC developmental abnormality, gastrointestinal ulceration, neuropathy,

CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid  
CC disorders e.g. hypothyroidism, brain damages, colitis, cone photo-  
CC transduction deficiency, neurological diseases, stroke, angiogenesis,  
CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,  
CC trachea, thymus, lymph node and muscular system, obesity, anorexia,  
CC growth abnormalities, and alleviation of precocious puberty. The  
CC sequences given in records AB086130-AB086184 represent novel human cdna's  
CC of the invention.

XX Sequence 3096 BP; 890 A; 619 C; 742 G; 845 T; 0 other;

SQ Query Match 17.5%; Score 161.6; DB 24; Length 3096;

Best Local Similarity 51.6%; Pred. No. 6.8e-41;

Matches 399; Conservative 0; Mismatches 369; Indels 6; Gaps 1;

QY 148 TTAACTTTTTCAGTCGCATCTTCTGGAGAGCCCAAGTGGAGAGGTTTCCTATCCCTGG 207

DB 73 TCTAGATTCTTCTCTAGAAATTTAGTTGGAGAAATTAACAGTCTGCTGGACATCCATGG 132

QY 208 CAGGTATCTCTGAAACAAAGCGAGAGCATATTTTGGAGGAGCATCTCTCACCACAG 267

DB 133 CAGGTCTCCCTAAATCAGATGAGCACCACCTTCTCTGGAGGAGCTTGATTCAAGAAGAT 192

QY 268 TGGGTGATCAGCGCGCTCAGTCGATTTGCAACAGA-----AACATTGTGCTACTTTG 321

DB 193 CGGGTTGTTACAGACGACACTGCTGGACAGCCTCAGTGAGAGCAGCTGAAGAATATA 252

QY 322 ATGTTACTGCTGGAGAGTATGACTTAAGCCAGACAGACCCAGGAGAGCAAACTCTCACT 381

DB 253 ACTGTGACTTCTGGGAGTACAGCCTCTTTCAGAGGAGTAAGCAAGACAGAAATATTCCT 312

QY 382 ATTTAACTGTCTATCATATACATACATTTCTCCACCAAGAAACCAATGACTATGATATT 441

DB 313 GTCTCAAAATTTATACCATCTCTGANTACACAGCCGTGAATATATGAGTCTGATATT 372

QY 442 GCCCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCACTTTTGGGGCCCATATGTCCT 501

DB 373 GCACGTCTGATATCTAAACACACAAAGTCAAGTTTGGAAATGCTTTCAGGCCAATCTGCTT 432

QY 502 CCAGAGCTGCGGGGACAAATTTGAGGCTGTTTATTTTGTACAACTGCAGGCTGGGGCCGC 561

DB 433 CCTGACAGGATGATAAGTTGAACAGGAATCTTTTGTCTTATCCAGTGGATGGGCAAG 492

QY 562 TTAACTGAAGTGGGCTCTCTCACAAAGTCTTTCAGGAGTGAATCTGCTATTTTGACC 621

DB 493 ATTTCCAAACATCAGAAATATTCAAATGTCTTACAAGAAATGAACTTCCCATCGAT 552

QY 622 TGGGAAGAGTGTGGCAGCTCTGTTTAAACACTAAAGAGGCCCATCAGTGGGAGACCTTT 681

DB 553 GACAGAGCGTGAATACTGTCTCAAGAGCATGAACCTCCCTCCCTGGAAGGACCATG 612

QY 682 CTTTCACACAGTTTCTCTGATGGAGGAGAGCAGCATGTCCAGGGAGATTCAGAGGTTCA 741

DB 613 CTGTGTGTGGCTTCCCTGATTTGGGAAATGGACGCTGCCAGGGGACTCTCGAGGACCA 672

QY 742 CTCATGTGCCGGAATAAGAAAGGGCGCTGGACTCTGGCTGTGTGCTTCTCTGGGTTTG 801

DB 673 CTGGTTGTAGAAAGTGGTGGAAATCTGGATTTCTTGTGGGATTAATCTCTCTGGTAGCT 732

QY 802 GGCTGTGTTCGAGGCTGGAGAAACAAATGTGAGGAAAGTGAATCAAGGATCCCTCGGATC 861

DB 733 GGTGTGTGGAGTTTCAGTTCCCTTAAGAAACACCATGTGAAGGCATCACTTGGCAATT 792

QY 862 TTCACAGACATTAGTAAAGTGTCTTCTCGGATTCACAGAACACATCAAACTGGT 915

DB 793 TTCTCCAAAGTGTCTGAGTTGATGATTTTATCTACTCAAAACCTGTTTCACAGAT 846

Search completed: March 25, 2003, 05:11:15

Job time : 285 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2003, 05:11:20 ; Search time 247 seconds  
(without alignments)  
2891.552 Million cell updates/sec

Title: US-09-735-713A-1  
Perfect score: 921  
Sequence: 1 atgagctctcaaaatgcttat.....acatccaaactggtaactaa 921

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 538826 seqs, 387737923 residues

Total number of hits satisfying chosen parameters: 1077652

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published\_Applications\_NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US05\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	920.2	99.9	921	9	US-09-735-713A-1
2	920.2	99.9	1568	9	US-09-735-713A-7
3	908.2	98.6	909	9	US-09-735-713A-3
4	771.2	83.7	1671	10	US-09-888-615-47
5	495	53.7	495	9	US-09-735-713A-5
6	114.6	12.4	2913	10	US-09-888-615-42
7	76.2	8.3	1327	9	US-09-978-295A-170
8	76.2	8.3	1327	9	US-09-978-697-170
9	76.2	8.3	1327	9	US-09-978-192A-170
10	76.2	8.3	1327	9	US-09-995-832A-170
11	76.2	8.3	1327	9	US-09-978-189-170
12	76.2	8.3	1327	9	US-09-978-608A-170
13	76.2	8.3	1327	9	US-09-978-191A-170
14	76.2	8.3	1327	9	US-09-978-403A-170
15	76.2	8.3	1327	9	US-09-978-564A-170
16	76.2	8.3	1327	9	US-09-978-585A-170
17	76.2	8.3	1327	9	US-10-017-081A-170
18	76.2	8.3	2409	10	US-09-888-615-54
19	76.2	8.3	3143	9	US-09-978-295A-168

ALIGNMENTS

RESULT 1

US-09-735-713A-1

; Sequence 1, Application US/09735713A

; Patent No. US20020165376A1

; GENERAL INFORMATION:

; APPLICANT: Walke, D. Wade

; APPLICANT: Turner, C. Alexander Jr.

; APPLICANT: Abuin, Alejandro

; APPLICANT: Friedrich, Glenn

; APPLICANT: Zambrowicz, Brian

; APPLICANT: Sands, Arthur T.

; TITLE OF INVENTION: No. US20020165376A1el Human Proteases and

; FILE REFERENCE: LEX-0108-USA

; CURRENT APPLICATION NUMBER: US/09/735,713A

; PRIOR FILING DATE: 2000-12-12

; PRIOR APPLICATION NUMBER: US 60/171,566

; PRIOR FILING DATE: 1999-12-22

; NUMBER OF SEQ ID NOS: 7

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 921

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-735-713A-1

Query Match 99.9% ; Score 920.2; DB 9; Length 921;

Best Local Similarity 100.0%; Pred. No. 2.4e-250; Mismatches 0; Indels 0; Gaps 0;

Matches 921; Conservative 0;

QY 1 ATGAGCTCTCAAAATGCTTATAAGCAGGACAAAGCTGATTTTACTACTAGGAATAGTCTTTT 60

Db 1 ATGAGCTCTCAAAATGCTTATAAGCAGGACAAAGCTGATTTTACTACTAGGAATAGTCTTTT 60

QY 61 TTGGAACRAGGTAATGTCARCTCTTTGCGTCCCAAGCTCCAGTTGTGGGAGACT 120

Db 61 TTGGAACRAGGTAATGTCARCTCTTTGCGTCCCAAGCTCCAGTTGTGGGAGACT 120

QY 121 CTGGTTAAGGTACAGCTTGGAAATTTATTTACATTTTTCAGTCGCAATCTTTGGAGAAGC 180

Db 121 CTGGTTAAGGTACAGCTTGGAAATTTATTTACATTTTTCAGTCGCAATCTTTGGAGAAGC 180

```
QY 181 CAAGTCGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAGAGGCAAGCATATT 240
Db 181 CAAGTCGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAGAGGCAAGCATATT 240
QY 241 TGTGGAGAGCATCGTCTCACCACAGTGGGTGATCAACGGCGCTCACTGCATTGCAAAAC 300
Db 241 TGTGGAGAGCATCGTCTCACCACAGTGGGTGATCAACGGCGCTCACTGCATTGCAAAAC 300
QY 301 AGAAACATTGTGCTACTTTTGAATGTTACTTCTGGAGAGTATGACTTTAAGCCACAGACAG 360
Db 301 AGAAACATTGTGCTACTTTTGAATGTTACTTCTGGAGAGTATGACTTTAAGCCACAGACAG 360
QY 361 CCAGGAGAGCAAACTCTCACTATTGAAACTGTCTATACATATGATGACTTAAAGCCACAGAC 420
Db 361 CCAGGAGAGCAAACTCTCACTATTGAAACTGTCTATACATATGATGACTTAAAGCCACAGAC 420
QY 421 AAACCAATGACATGATATGCGCTTTTGAAGATGCTGGAGCCTTCCAAATTGGCCAC 480
Db 421 AAACCAATGACATGATATGCGCTTTTGAAGATGCTGGAGCCTTCCAAATTGGCCAC 480
QY 481 TTTGTGGGGCCCATATCTCTCCAGAGCTCGGGAGCAATTTGAGGCTGGTTTATTGT 540
Db 481 TTTGTGGGGCCCATATCTCTCCAGAGCTCGGGAGCAATTTGAGGCTGGTTTATTGT 540
QY 541 ACAACTCGAGCTGGGGCCGCTTAACCTGAAGTGGCGTCTCTCACAAGTCTTGCAGGAA 600
Db 541 ACAACTCGAGCTGGGGCCGCTTAACCTGAAGTGGCGTCTCTCACAAGTCTTGCAGGAA 600
QY 601 GTGAATCTGCTATTTGACCTGGAGAGTGTGTGCAGCTCTGTAAACATAAGAGG 660
Db 601 GTGAATCTGCTATTTGACCTGGAGAGTGTGTGCAGCTCTGTAAACATAAGAGG 660
QY 661 CCCATCAGTGGGAGACCTTTCTTTGCACAGGTTTTCCTGATGAGGAGAGACGCATGT 720
Db 661 CCCATCAGTGGGAGACCTTTCTTTGCACAGGTTTTCCTGATGAGGAGAGACGCATGT 720
QY 721 CAGGAGATTACAGAGTTCACTCATGTGCGGGAATAAGAAAGGGGCTCGACTCTGGCT 780
Db 721 CAGGAGATTACAGAGTTCACTCATGTGCGGGAATAAGAAAGGGGCTCGACTCTGGCT 780
QY 781 GGTGTGACTTCCCTGGGCTTTGGGCTGTGTCGAGGCTGGAGAACAAATGTGAGGAAAAGT 840
Db 781 GGTGTGACTTCCCTGGGCTTTGGGCTGTGTCGAGGCTGGAGAACAAATGTGAGGAAAAGT 840
QY 841 GATCAAGATCCCTCGGATCTTACACAGACATTAGTAAAGTGTCTTCTGGATCCACGAA 900
Db 841 GATCAAGATCCCTCGGATCTTACACAGACATTAGTAAAGTGTCTTCTGGATCCACGAA 900
QY 901 CACATCCAAACTGGTAACTAA 921
Db 901 CACATCCAAACTGGTAACTAA 921
```

RESULT 2

```
US-09-735-713a-7
; Sequence 7, Application US/09735713a
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abulin, Alejandro
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020165376A1el Human Proteases and
; FILE REFERENCE: Polynucleotides Encoding the Same
; CURRENT APPLICATION NUMBER: US/09/735,713A
; PRIOR FILING DATE: 2000-12-12
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 7
; LENGTH: 1568
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-713a-7
```

```
Query Match 99.9%; Score 920.2; DB 9; Length 1568;
Best Local Similarity 100.0%; Pred. No. 3.3e-290;
Matches 921; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGAGTCTCAAAATGCTTATAAGCAGGAACAAGCTGATTTTACTACTAGGAATAGTCTTT 60
Db 245 ATGAGTCTCAAAATGCTTATAAGCAGGAACAAGCTGATTTTACTACTAGGAATAGTCTTT 304
QY 61 TTTGAACRAGGTAAATCTGCARCTCTTTCTCCTCCCAAAAGCTCCACAGTCTTGGGCAGAGT 120
Db 305 TTTGAACRAGGTAAATCTGCARCTCTTTCTCCTCCCAAAAGCTCCACAGTCTTGGGCAGAGT 364
QY 121 CTGTTTAAGGTACAGCCTTGGNAATTAATTTAACATATTTTCAGTCGCAATTTTGGAGGAAGC 180
Db 365 CTGTTTAAGGTACAGCCTTGGNAATTAATTTAACATATTTTCAGTCGCAATTTTGGAGGAAGC 424
QY 181 CAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAGGCAAGCATATT 240
Db 425 CAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAGGCAAGCATATT 484
QY 241 TGTGGAGGAAGCATCGTCTCACCACAGTGGGTGATCAACGGCGCTCACTGCATTGCAAAAC 300
Db 485 TGTGGAGGAAGCATCGTCTCACCACAGTGGGTGATCAACGGCGCTCACTGCATTGCAAAAC 544
QY 301 AGAAACATTGTGCTACTTTTGAATGTTTACTGCTGGAGAGTATGACTTTAAAGCCACAGACAG 360
Db 545 AGAAACATTGTGCTACTTTTGAATGTTTACTGCTGGAGAGTATGACTTTAAAGCCACAGACAG 604
QY 361 CCAGGAGAGCAAACTCTCACTATTTGAACATGTCATCATATCCACATTTTCTCCACCAAG 420
Db 605 CCAGGAGAGCAAACTCTCACTATTTGAACATGTCATCATATCCACATTTTCTCCACCAAG 664
QY 421 AAACCAATGACATGATATTTGCCCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCAC 480
Db 665 AAACCAATGACATGATATTTGCCCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCAC 724
QY 481 TTTGTGGGGCCCATATGTCTTCCAGAGCTCGGGAGCAATTTGAGGCTGGTTTATTGT 540
Db 725 TTTGTGGGGCCCATATGTCTTCCAGAGCTCGGGAGCAATTTGAGGCTGGTTTATTGT 784
QY 541 ACAACTGCAGGCTGGGGCCGCTTAACTGAAGTGGGCTCTCTCACAAGTCTTGCAGGAA 600
Db 785 ACAACTGCAGGCTGGGGCCGCTTAACTGAAGTGGGCTCTCTCACAAGTCTTGCAGGAA 844
QY 601 GTGAATCTGCCTATTTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTAAACACTAAAGAGG 660
Db 845 GTGAATCTGCCTATTTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTAAACACTAAAGAGG 904
QY 661 CCCATCAGTGGGAGACCTTTCTTTGCACAGGTTTTCCTGATGAGGAGAGACGCATGT 720
Db 905 CCCATCAGTGGGAGACCTTTCTTTGCACAGGTTTTCCTGATGAGGAGAGACGCATGT 964
QY 721 CAGGAGATTACAGAGTTCACTCATGTGCCGGAATAAGAAAGGGGCTCGACTCTGGCT 780
Db 965 CAGGAGATTACAGAGTTCACTCATGTGCCGGAATAAGAAAGGGGCTCGACTCTGGCT 1024
QY 781 GGTGTGACTTCCCTGGGCTTTGGGCTGTGTCGAGGCTGGAGAACAAATGTGAGGAAAAGT 840
Db 1025 GGTGTGACTTCCCTGGGCTTTGGGCTGTGTCGAGGCTGGAGAACAAATGTGAGGAAAAGT 1084
QY 841 GATCAAGGATCCCTCGGATCTTACACAGACATTAGTAAAGTGTCTTCTGGATCCACGAA 900
Db 1085 GATCAAGGATCCCTCGGATCTTACACAGACATTAGTAAAGTGTCTTCTGGATCCACGAA 1144
QY 901 CACATCCAAACTGGTAACTAA 921
Db 1145 CACATCCAAACTGGTAACTAA 1165
```

```
RESULT 3
US-09-735-713A-3
; Sequence 3, Application US/09735713A
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abuin, Alejandro
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0108-USA
; CURRENT APPLICATION NUMBER: US/09/735,713A
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,566
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 909
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-713A-3

Query Match      98.6%; Score 908.2; DB 9; Length 909;
Best Local Similarity 100.0%; Pred. No. 2e-286;
Matches 909; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 ATGCTTAAACGACGACAAAGCTGATTTTACTAGTAAGTACTGTTTGAACRAGGT 72
DB 1 ATGCTTAAACGACGACAAAGCTGATTTTACTAGTAAGTACTGTTTGAACRAGGT 60
QY 73 AAATCTGCARCTCTTTTCGCTCCCCAAAGCTCCCAAGTGTGGGCGAGAGTCTGGTTAAAGTA 132
DB 61 AAATCTGCARCTCTTTTCGCTCCCCAAAGCTCCCAAGTGTGGGCGAGAGTCTGGTTAAAGTA 120
QY 133 CAGCCTTGGAAATTTTAAACATTTTCAGTCGCAATTTCTGGAGAGCCCAAGTGGAGAG 192
DB 121 CAGCCTTGGAAATTTTAAACATTTTTCAGTCGCAATTTCTGGAGAGCCCAAGTGGAGAG 180
QY 193 GTTCTCTATCCCTGGCAGGTATCTCTGAACAAAGGCGAGAGCATATTTGTGGAGGAGC 252
DB 181 GTTCTCTATCCCTGGCAGGTATCTCTGAACAAAGGCGAGAGCATATTTGTGGAGGAGC 240
QY 253 ATCGTCTACACAGTGGGTGATCAGCGGCTCAGTGCATTTGCAAAACAGAAACATTTGTG 312
DB 241 ATCGTCTACACAGTGGGTGATCAGCGGCTCAGTGCATTTGCAAAACAGAAACATTTGTG 300
QY 313 TCTACTTTGAATGTTACTGTGGAGATGATGACTTAAGCCAGACAGACCCAGGAGAGCAA 372
DB 301 TCTACTTTGAATGTTACTGTGGAGATGATGACTTAAGCCAGACAGACCCAGGAGAGCAA 360
QY 373 ACTCTCACTATTGAACCTGTGCATCATACATCCACATTTCTCCACCAAGAACCAATGAC 432
DB 361 ACTCTCACTATTGAACCTGTGCATCATACATCCACATTTCTCCACCAAGAACCAATGAC 420
QY 433 TATGATATTGCCCTTTTGAAGATGGCTGGAGCCCTTCCAATTTGGCCACTTTGTGGGGCCC 492
DB 421 TATGATATTGCCCTTTTGAAGATGGCTGGAGCCCTTCCAATTTGGCCACTTTGTGGGGCCC 480
QY 493 ATATGCTTTCAGAGCTCGGGAGCAATTTGAGGCTGGTGTATTTTGTACAACTGCGAGGC 552
DB 481 ATATGCTTTCAGAGCTCGGGAGCAATTTGAGGCTGGTGTATTTTGTACAACTGCGAGGC 540
QY 553 TGGGGCCGCTTAACTGAGGTGGCGTCTCTCAACAGTCTTTCAGGAGTGAATCTGCCT 612
DB 541 TGGGGCCGCTTAACTGAGGTGGCGTCTCTCAACAGTCTTTCAGGAGTGAATCTGCCT 600
QY 613 ATTTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTGTAACACTTAAAGAGGCCCATCAGTGGG 672
```

```
Db 601 ATTTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTGTAACACTTAAAGAGGCCCATCAGTGGG 660
QY 673 AAGACCTTTCTTTTCACAGAGTGTTCCTGATGAGGGAGAGAGCGCATGTCAGGGAGATTCA 732
DB 661 AAGACCTTTCTTTTCACAGAGTGTTCCTGATGAGGGAGAGAGCGCATGTCAGGGAGATTCA 720
QY 733 GGAGGTTTCACTCATGTGCCGAATAAGAAAGGCGCTGGACTCTGGCTGGTGTGACTTCC 792
DB 721 GGAGGTTTCACTCATGTGCCGAATAAGAAAGGCGCTGGACTCTGGCTGGTGTGACTTCC 780
QY 793 TGGGTTTGGGCTGTGTGCGAGGCTGGAGAAACAATGTGAGGAAAGTGTATCAAGGATCC 852
DB 781 TGGGTTTGGGCTGTGTGCGAGGCTGGAGAAACAATGTGAGGAAAGTGTATCAAGGATCC 840
QY 853 CCTGGGATCTTTCACAGACATTTAGTAAAGTGTCTTCTCGGATCCAGGAACATCCAAACT 912
DB 841 CCTGGGATCTTTCACAGACATTTAGTAAAGTGTCTTCTCGGATCCAGGAACATCCAAACT 900
QY 913 GGTAACATAA 921
DB 901 GGTAACATAA 909

RESULT 4
US-09-888-615-47
; Sequence 47, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARVDCZAK, GLEN
; APPLICANT: MANNING, GERRARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-888-615-47

Query Match      83.7%; Score 771.2; DB 10; Length 1671;
Best Local Similarity 99.6%; Pred. No. 1.8e-241;
Matches 781; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 ATGAGTCTCAAAATGCTTTATAGCAGGAAACAAGCTGATTTTACTACTAGGAATAGTCTTT 60
DB 1 ATGAGTCTCAAAATGCTTTATAGCAGGAAACAAGCTGATTTTACTACTAGGAATAGTCTTT 60
QY 61 TTTGAACRAGGTAATCTGCACTCTTTCGCTCCCCAAAGCTCCAGTGTGGGCGAGAT 120
DB 61 TTTGAACRAGGTAATCTGCACTCTTTCGCTCCCCAAAGCTCCAGTGTGGGCGAGAT 120
QY 121 CTGGTTAAGGTACAGCCTTGGAAATTTTAAACATTTTTCAGTCGATCTTTGGAGGAAGC 180
DB 121 CTGGTTAAGGTACAGCCTTGGAAATTTTAAACATTTTTCAGTCGATCTTTGGAGGAAGC 180
QY 181 CAAGTGGAGAAAGGTTTCTATCCCTGGCAGGTATCTCTGAAACAAAGCAGAGCATATT 240
DB 181 CAAGTGGAGAAAGGTTTCTATCCCTGGCAGGTATCTCTGAAACAAAGCAGAGCATATT 240
QY 241 TGTGGAGAAACATCGTCTCACACAGTGGGTGATCAGCGGGCTCAGTGCATTTGCAAAAC 300
DB 241 TGTGGAGAAACATCGTCTCACACAGTGGGTGATCAGCGGGCTCAGTGCATTTGCAAAAC 300
```

```
QY 301 AGAACAATTGTGTACTTGTGAATGTACTGTGGAGTATGACTTAAAGCCAGACAGAC 360
Db 301 AGAACAATTGTGTACTTGTGAATGTACTGTGGAGTATGACTTAAAGCCAGACAGAC 360
QY 361 CCAGGAGACAAACTCTCACTTGAACCTGTCATCATACATCCACATTTCTCCACCAG 420
Db 361 CCAGGAGACAAACTCTCACTTGAACCTGTCATCATACATCCACATTTCTCCACCAG 420
QY 421 AAACCAATGGACTATGATATGCGCTTTTGAAGATGGCTGGAGCCCTTCCAAATTTGGCCAC 480
Db 421 AAACCAATGGACTATGATATGCGCTTTTGAAGATGGCTGGAGCCCTTCCAAATTTGGCCAC 480
QY 481 TTTGTGGGCCCATATGCTCTCCAGAGCTGCGGGAGCAATTTGAGCGTGGTTTTATTTGT 540
Db 481 TTTGTGGGCCCATATGCTCTCCAGAGCTGCGGGAGCAATTTGAGCGTGGTTTTATTTGT 540
QY 541 ACAACTGCAGCTGGGGCGGCTTAAGTGAAGTGGCTCTCTCAAGTCTTGAGGAA 600
Db 541 ACAACTGCAGCTGGGGCGGCTTAAGTGAAGTGGCTCTCTCAAGTCTTGAGGAA 600
QY 601 GTGAATCTCCCTATTTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTTAACACTTAAAGAGG 660
Db 601 GTGAATCTCCCTATTTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTTAACACTTAAAGAGG 660
QY 661 CCATCAGTGGGAAGACCTTTCTTGCCAGAGTGTTCCTGATGGAGGGAGACGCATGT 720
Db 661 CCATCAGTGGGAAGACCTTTCTTGCCAGAGTGTTCCTGATGGAGGGAGACGCATGT 720
QY 721 CAGGAGATTCAGGAGTTCACTCATGTGCCGGAATAAGAAAGGGGCTT-GGACTCTGCC 779
Db 721 CAGGAGATTCAGGAGTTCACTCATGTGCCGGAATAAGAAAGGGGCTTGGGACTCTGCC 780
QY 780 TGGT 783
Db 781 TGGT 784

RESULT 5
US-09-735-713A-5
; Sequence 5, Application US/09735713A
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abulin, Alejandro
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020165376A1el Human Proteases and
; FILE REFERENCE: Polynucleotides Encoding the Same
; CURRENT FILING DATE: US/09/735,713A
; PRIOR FILING DATE: 2000-12-12
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 7
; SEQ ID NO 5
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-713A-5

Query Match 53.7%; Score 495; DB 9; Length 495;
Best Local Similarity 100.0%; Pred. No. 1.7e-151;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 427 ATGGACTATGATATGCCCTTTTGAAGATGGCTGGAGCCCTTCCAAATTTGGCCACTTTGTG 486
Db 1 ATGGACTATGATATGCCCTTTTGAAGATGGCTGGAGCCCTTCCAAATTTGGCCACTTTGTG 60
QY 487 GGGCCCATATGCTTCCAGAGCTGCGGAGCAATTTGAGGCTGGTTTTATTTGTACAAC 546
```

```
Db 61 GGGCCCATATGCTTCCAGAGCTGCGGAGCAATTTGAGGCTGGTTTTATTTGTACAAC 120
QY 547 GCAGGCTGGGGCGGCTTAACTGAAGTGGCGCTCTCTCAACAAGTCTTGCAGGAAGTGAAT 606
Db 121 GCAGGCTGGGGCGGCTTAACTGAAGTGGCGCTCTCTCAACAAGTCTTGCAGGAAGTGAAT 180
QY 607 CTGCTTATTTTGAACCTGGGAAGAGTGTGTGGCAGCTCTCTTAACACTAAAGAGGGCCCATC 666
Db 181 CTGCTTATTTTGAACCTGGGAAGAGTGTGTGGCAGCTCTCTTAACACTAAAGAGGGCCCATC 240
QY 667 AGTGGGAAGACCTTTCTTTTGCACAGGTTTCTGATGGAGGAGAGACGCATGTCAAGGA 726
Db 241 AGTGGGAAGACCTTTCTTTTGCACAGGTTTCTGATGGAGGAGAGACGCATGTCAAGGA 300
QY 727 GATTGAGGAGTTCACTCATGTGCCGGAATAAGAAAGGGGCTTGGAGTCTGGCTGGTGTG 786
Db 301 GATTGAGGAGTTCACTCATGTGCCGGAATAAGAAAGGGGCTTGGAGTCTGGCTGGTGTG 360
QY 787 ACTTCTCTGGGTTTGGGCTGTGTGCGAGCTGGAGAAACAATGTGAGGAAAAAGTGTATCAA 846
Db 361 ACTTCTCTGGGTTTGGGCTGTGTGCGAGCTGGAGAAACAATGTGAGGAAAAAGTGTATCAA 420
QY 847 GGATCCCCGGGATCTTTCACAGACATTTAGTAAAGTCTTCTTGGATCCACGACACATC 906
Db 421 GGATCCCCGGGATCTTTCACAGACATTTAGTAAAGTCTTCTTGGATCCACGACACATC 480
QY 907 CAAACTGGTAACTAA 921
Db 481 CAAACTGGTAACTAA 495

RESULT 6
US-09-888-615-42
; Sequence 42, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARVDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 42
; LENGTH: 2913
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-888-615-42

Query Match 12.4%; Score 114.6; DB 10; Length 2913;
Best Local Similarity 52.4%; Pred. No. 1.3e-26;
Matches 252; Conservative 0; Mismatches 229; Indels 0; Gaps 0;
```

```
QY 435 TGATATTTGCCCTTTTGAAGATGGCTGGAGCCCTTCCAATTTGGCCACTTTTGTGGGCCCAT 494
Db 9 TGATATTTGCCCTTTTGAAGATGGCTGGAGCCCTTCCAATTTGGCCACTTTTGTTCAGCCAA 68
QY 495 ATGTCTTCCAGAGCTGCGGAGCAATTTGAGGCTGGTTTTATTTGTACAACTGCAGGCTG 554
Db 69 CTGTCTTCTGACAGGATGATAAGTTGAACCAAGAAATCTTTGCTTATCAAGTGGATG 128
QY 555 GGGCGCGCTTAACTGAAGTGGCGCTCTCTCAACAAGTCTTTCAGGAAGTGAATCTGCCTAT 614
Db 129 GGGCAAGATTTCCAAAACATCAGATATTCAAATGTCTTACAAAGAAATGGAATCTCCCAT 188
QY 615 TTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTTAACTAAAGAGGCCCATCAGTGGGAA 674
```

Db 189 CATGGATGACAGCGCTGTAATACTGTGCTCAAGACCATGAACCTCCCTCCCTGGGAAG 248  
QY 675 GACCTTTCTTTGACAGGTTTCCCTGATGGAGGAGAGAGCATGTGACGGGAGATTCCAGG 734  
Db 249 GACCATGCTGTGCTGGCTCCCTGATTTGGGAATGGAGCGCTGCCAGGGGACACTCTGG 308  
QY 735 AGTTTCACATCATGTGCGCGGAATAAGAAAGGCGCTGGACTCTGGCTGGTGTGACTTCCCTG 794  
Db 309 AGGACCATGCTTTGTAGAGAGGTGGTGAATCTGGATCTGTGCTGGGATACATTCCTG 368  
QY 795 GGGTTTGGGCTGTGCTGGAGGTGGAGAAACAATGTGAGGAAAGATGATCAAGGATCCCC 854  
Db 369 GGTAGTGGTGTGCTGGAGGTTCAGTTCCCGTAAGAAACAACCATGTGAAAGGCATCACT 428  
QY 855 TGGGATCTTCACAGACATTAAGTAAAGTCTTCCCTGGATCCAGGACACATCCAAACTGG 914  
Db 429 TGGCATTTTCTCCAAAGTCTGAGTGTGATGATTTTATCACTCAAAACCTGTTACAGG 488  
QY 915 T 915  
Db 489 T 489

RESULT 7  
US-09-978-295A-170  
; Sequence 170, Application US/09978295A  
; Patent No. US20020156006A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan  
; APPLICANT: Ferrara, Napoleon  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Kuo, Sophia S.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James;  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Shelton, David L.  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2630PIC11  
; CURRENT APPLICATION NUMBER: US/09/978,295A  
; CURRENT FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: 09/918585  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/064249  
; PRIOR FILING DATE: 1997-11-03  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066364  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: 60/077450  
; PRIOR FILING DATE: 1998-03-10  
; PRIOR APPLICATION NUMBER: 60/077632  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077641  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077649  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077791  
; PRIOR FILING DATE: 1998-03-12  
; PRIOR APPLICATION NUMBER: 60/078004  
; PRIOR FILING DATE: 1998-03-13  
; PRIOR APPLICATION NUMBER: 60/078886  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078936  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078939  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079294  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: 60/079656  
; PRIOR FILING DATE: 1998-03-26  
; PRIOR APPLICATION NUMBER: 60/079664  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079689  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079663  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079728  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079786  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079920  
; PRIOR FILING DATE: 1998-03-30  
; PRIOR APPLICATION NUMBER: 60/079923  
; PRIOR FILING DATE: 1998-03-30  
; PRIOR APPLICATION NUMBER: 60/080105  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080107  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080165  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080194  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080327  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080328  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080333  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080334  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/081070  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081049  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081071  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081195  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081203  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081229  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081955  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081817  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081819  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081952  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081838



APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630PIC27  
CURRENT APPLICATION NUMBER: US/09/978,697  
CURRENT FILING DATE: 2001-10-16  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366

;  
; PRIOR FILING DATE: 1998-05-05  
; PRIOR APPLICATION NUMBER: 60/084414  
; PRIOR FILING DATE: 1998-05-06  
; PRIOR APPLICATION NUMBER: 60/084441  
; PRIOR FILING DATE: 1998-05-06  
; PRIOR APPLICATION NUMBER: 60/084637  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084639  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084640  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084598  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084627  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084643  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/085339  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/085338  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/085323  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/085582  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085700  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085689  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085579  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085580  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085573  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085704  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 8.3%; Score 76.2; DB 9; Length 1327;  
Best Local Similarity 47.2%; Pred. No. 2.8e-14;  
Matches 311; Conservative 0; Mismatches 333; Indels 15; Gaps 2;  
  
QY 157 TTCAGTGCATCTTTGGAGGAAGCCCAAGTGGAGAGGCTTCTATCTCTGGCAGGTATCT 216  
DB 531 TCCAGCCGATTTGGTGGAGCTGTCTCCCGAGGGTGAGTGGCCATGGCAGGCCAGC 590  
  
QY 217 CTGAACAAAGCAGACATATTTGTGGAGAACATCGTCTCACCACAGTGGGTGATC 276  
DB 591 CTCCAGGTTGGGGTTCGACACATCTGTGGGGGGCCCTCATCGCTGACCGCTGGGTGATA 650  
  
QY 277 ACGGGCGCTCACTGATTCGAACAGAACATTTGTCTACTTGAATGTTACTGCTCGA 336  
DB 651 ACAGTGGCCACTGTCTCCAGGAGGACAGCATGGCCTCCACGGTCTGTGGACCGGTGTC 710  
  
QY 337 GAGTATGACTTAAAGCCAGACAGACCCAGGA---GAGCAAACTCTCACTATTGAAGTCTC 393  
DB 711 CTGGGCAAGGTGTGCGACAGACTCGCGCTGGAGAGGTGTCTTCAAGGTGAGCCGC 770  
  
QY 394 ATCATATCCACATTTCTCCACCAAGAACAAATGGACTATGATPATTTGCCCTTTTGAAG 453  
DB 771 CTGCTCTGCAACCCGTACCAAGAGGACAGCCATGACTAGCAGGTGGCGCTGCTGCAG 830  
  
QY 454 ATGGCTGAGCCTTCCATTTGGCCACATTTGTGGGGCCCATATGCTTCCAGAGTGGCG 513  
DB 831 CTCGACCAACCCCGGTGGTGGCTGGCCCGCGCTGCGCCCGCTGCTGCGCGCGCTCC 890  
  
QY 514 GAGCAATTGAGCGTGGTTTATTTGTACAACTGCAGCTGGCGCGCTTAACCTAAGGT 573  
DB 891 CACTTCTCGAGCCCGGCTGCACCTGCTGGATTACGGGCTGGGGCGCTTGGCGGAGGC 950

QY 574 GGCGTCTCTCACAGTCTTGCAGGAAGTGAATCTGCCTATTTTGAACCTGGGAAGAGTGT 633  
DB 951 GGCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTGATCCACAGGACCTGTGC 1010  
  
QY 634 GTGCAGCTCTGTTAACACTAAAGAGGCCCATCAGTGGGAAGACCTTTCTTTGCACAGT 693  
DB 1011 AGCGAGGCTATCGCTACCGTGTACGCCACGCA-----TGCTGTGTGCCGC 1058  
  
QY 694 TTTCTGATGGAGGAGACGCGATGTGAGGAGATTTCAGGAGGTTCACTATGTGCCGG 753  
DB 1059 TACCGAAGGCAAGAGGATGCTGTCAGGGTGACTCAGTGGTCCCTGTTGTCAGAG 1118  
  
QY 754 AATAAGAAAGGGCCTGGACTCTGGCTGGTGTGACTTCTCTGGGGTTGGGCTGTGGTGC 812  
DB 1119 GCAGTCACTGGCGCGTGTCTCTGGCGGGGCTGGTCACTGGGGCTGTGGCGG 1177

RESULT 9  
US-09-978-192A-170  
; Sequence 170, Application US/09978192A  
; Patent No. US20020177553A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan  
; APPLICANT: Ferrara, Napoleon  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gottard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Kuo, Sophia S.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James;  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Shelton, David L.  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2630P1C9  
; CURRENT APPLICATION NUMBER: US/09/978,192A  
; CURRENT FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: 09/918585  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/064249  
; PRIOR FILING DATE: 1997-11-03  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066364  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: 60/077450  
; PRIOR FILING DATE: 1998-03-10  
; PRIOR APPLICATION NUMBER: 60/077632  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077641  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077649  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077791



```

: APPLICANT: Napier, Mary A.
: APPLICANT: Pan, James;
: APPLICANT: Paoni, Nicholas F.
: APPLICANT: Roy, Margaret Ann
: APPLICANT: Shelton, David L.
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Williams, P. Mickey
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
: TITLE OF INVENTION: Acids Encoding the Same
: FILE REFERENCE: P2630PIC63
: CURRENT APPLICATION NUMBER: US/09/999,832A
: CURRENT FILING DATE: 2001-10-24
: PRIOR APPLICATION NUMBER: 09/918585
: PRIOR FILING DATE: 2001-07-30
: PRIOR APPLICATION NUMBER: 60/062250
: PRIOR FILING DATE: 1997-10-17
: PRIOR APPLICATION NUMBER: 60/064249
: PRIOR FILING DATE: 1997-11-03
: PRIOR APPLICATION NUMBER: 60/065311
: PRIOR FILING DATE: 1997-11-13
: PRIOR APPLICATION NUMBER: 60/066364
: PRIOR FILING DATE: 1997-11-21
: PRIOR APPLICATION NUMBER: 60/077450
: PRIOR FILING DATE: 1998-03-10
: PRIOR APPLICATION NUMBER: 60/077632
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077641
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077649
: PRIOR FILING DATE: 1998-03-11
: PRIOR APPLICATION NUMBER: 60/077791
: PRIOR FILING DATE: 1998-03-12
: PRIOR APPLICATION NUMBER: 60/078004
: PRIOR FILING DATE: 1998-03-13
: PRIOR APPLICATION NUMBER: 60/078886
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078936
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078910
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/078939
: PRIOR FILING DATE: 1998-03-20
: PRIOR APPLICATION NUMBER: 60/079294
: PRIOR FILING DATE: 1998-03-25
: PRIOR APPLICATION NUMBER: 60/079656
: PRIOR FILING DATE: 1998-03-26
: PRIOR APPLICATION NUMBER: 60/079664
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079689
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079663
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079728
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079786
: PRIOR FILING DATE: 1998-03-27
: PRIOR APPLICATION NUMBER: 60/079920
: PRIOR FILING DATE: 1998-03-30
: PRIOR APPLICATION NUMBER: 60/079923
: PRIOR FILING DATE: 1998-03-30
: PRIOR APPLICATION NUMBER: 60/080105
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080107
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080165
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080194
: PRIOR FILING DATE: 1998-03-31
: PRIOR APPLICATION NUMBER: 60/080327
: PRIOR FILING DATE: 1998-04-01
: PRIOR APPLICATION NUMBER: 60/080328

```

7	PRIOR FILING DATE: 1998-03-27	
7	PRIOR APPLICATION NUMBER: 60/079689	
7	PRIOR FILING DATE: 1998-03-27	
7	PRIOR APPLICATION NUMBER: 60/079663	
7	PRIOR FILING DATE: 1998-03-27	
7	PRIOR APPLICATION NUMBER: 60/079728	
7	PRIOR FILING DATE: 1998-03-27	
7	PRIOR APPLICATION NUMBER: 60/079786	
7	PRIOR FILING DATE: 1998-03-27	
7	PRIOR APPLICATION NUMBER: 60/079920	
7	PRIOR FILING DATE: 1998-03-30	
7	PRIOR APPLICATION NUMBER: 60/079923	
7	PRIOR FILING DATE: 1998-03-30	
7	PRIOR APPLICATION NUMBER: 60/080105	
7	PRIOR FILING DATE: 1998-03-31	
7	PRIOR APPLICATION NUMBER: 60/080107	
7	PRIOR FILING DATE: 1998-03-31	
7	PRIOR APPLICATION NUMBER: 60/080165	
7	PRIOR FILING DATE: 1998-03-31	
7	PRIOR APPLICATION NUMBER: 60/080194	
7	PRIOR FILING DATE: 1998-03-31	
7	PRIOR APPLICATION NUMBER: 60/080327	
7	PRIOR FILING DATE: 1998-04-01	
7	PRIOR APPLICATION NUMBER: 60/080328	







; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2630P1C22  
; CURRENT APPLICATION NUMBER: US/09/978, 608A  
; CURRENT FILING DATE: 2001-10-16  
; NUMBER OF SEQ ID NOS: 624  
; Prior Application removed - See File Wrapper or Palm  
; SEQ ID NO 170  
; LENGTH: 1327  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-978-608A-170

Query Match 8.3%; Score 76.2; DB 9; Length 1327;  
Best Local Similarity 47.2%; Pred. No. 2.8e-14;  
Matches 311; Conservative 0; Mismatches 333; Indels 15; Gaps 2;  
Qy 157 TTCAGTCGCATTCCTGTGAGGAAGCAAGTGGAGAAAGGTTCCCTATCCCTGGCAGGTATCT 216  
Db 531 TCCAGCGCATGTGTGGAGCTGTCTCTCCGAGGCTGAGTGGCCATGGCAGGCCAGC 590  
Qy 217 CTGAACAAGGCAAGATATTTGTGGAGGAAGCATCTCTCACACAGTGGGTGATC 276  
Db 591 CTCAGGTTTCGGGTCGACACATCTGTGGGGGGCCCTCATCGCTGACCGTGGGTGATA 650  
Qy 277 ACGCGCTCCTCACTGCAATGCAACAGAACATGTGTCTACTTTGATGTACTGCTGGA 336  
Db 651 ACAGTCCCACTCTTCCAGGAGGACAGCATGCGCTCCACGGTGTGTGGACCGTTC 710  
Qy 337 GAGTATGACTTAAAGCCAGACAGACCCAGGA---GAGCAAACTCTCACTATTGAAACTGTC 393  
Db 711 CTGGCAAGGTGTGCAGAACTCGCGCTGGCTGGAGAGGTCTCTCAAGGTGACGCCG 770  
Qy 394 ATCATATCATCATTTCTCCACCAAGAACCAATGAGTACATATGATGTCCTTTTGAAG 453  
Db 771 CTGCTCTGCACCGGTACCAAGAGAGAGACAGCCATGACTACGACGTGGCGCTGCTGCAG 830  
Qy 454 ATGCTGGAGCCTTCCAAATTTGGCCACTTTGTGGGGCCCATATGTCTTCAGAGAGCTCGG 513  
Db 831 CTCGACCAACCGGTGGTGGCTCGGCCCGCGTGGCCCGCTGCTGCGCCGCGCTCC 890  
Qy 514 GAGCAATTTGAGGTGTTTTTATTTGACAACTGACAGCTGGGGCCCTTAACTGAAGGT 573  
Db 891 CACTTCTTCGAGCCGCGCTGCACTGTGATTTACGGCTGGGGCGCTTTCGCGAGGGC 950  
Qy 574 GGCCTCTCTCACAGTCTTGCAGGAAGTGAATCTGCTTATTTGACCTGGGAGAGTGT 633  
Db 951 GGCCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTGATCCACAGGACCTGTGC 1010  
Qy 634 GTGCAGCTCTGTTAAACACTAAAGAGGCCCATCAGTGGGAAGACCTTTCTTTTGCACAGGT 693  
Db 1011 AGCAGGCTTATCGCTACCAAGGTGACCCACGCA-----TGCTGTGTGCGCGC 1058  
Qy 694 TTTCCCTGATGAGGAGAGACGATGTCAGGAGATTCAGAGGTTCACTCATGTGCGCG 753  
Db 1059 TACGCAAGGCAAGAGAGATGCTGTGAGGGTGACTCAGGTGCTGCGCTGGTGTGCAAG 1118  
Qy 754 AATAAGAAAGGGCCTGGACTCTGGCTGGTGTGACTTCCTTCCCTGGGCTTGGGCTGTGTCG 812  
Db 1119 GCATCAGTGGCGCTGGTCTTCCCTGGCGGGCTGTGTCAGCTGGGCGCTGTGGCGC 1177

RESULT 13

US-09-978-191A-170  
; Sequence 170, Application US/09978191A  
; Publication No. US20030050239A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc

; APPLICANT: Baton, Dan  
; APPLICANT: Ferrara, Napoleon  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Kuo, Sophia S.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Shelton, David L.  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2630P1C4  
; CURRENT APPLICATION NUMBER: US/09/978, 191A  
; CURRENT FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: 09/918585  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/064249  
; PRIOR FILING DATE: 1997-11-03  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066364  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: 60/077450  
; PRIOR FILING DATE: 1998-03-10  
; PRIOR APPLICATION NUMBER: 60/077632  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077641  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077649  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077791  
; PRIOR FILING DATE: 1998-03-12  
; PRIOR APPLICATION NUMBER: 60/078004  
; PRIOR FILING DATE: 1998-03-13  
; PRIOR APPLICATION NUMBER: 60/078886  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078936  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078939  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079294  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: 60/079656  
; PRIOR FILING DATE: 1998-03-26  
; PRIOR APPLICATION NUMBER: 60/079664  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079689  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079663  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079728  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079786  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079920



```
Db 831 CTCGACCCCGGTGGTGGCTCGCGCGCGTGGCGCCCGTCTGCGTGGCGCGCTCC 890
QY 514 GAGCAATTGAGCTGGTTTATTATTTGACAACTGCAGCTGGCGCGCTTAACTGAAGT 573
Db 891 CACTTCTTCGAGCGCGCTGCACCTGCTGGATTACGGCTGGGGCGCTTGGCGAGGGC 950
QY 574 GCGTCCTCTCAACAAGTCTTGAGGAAGTGAATCTGCTTATTTTACCTGGGAAGAGTGT 633
Db 951 GCGCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTGATCCACAGACCTGTC 1010
QY 634 GTGCGAGCTCTTTACACTAAAGAGGCCATCACTGGGAAGACCTTTCTTTGACACAGT 693
Db 1011 AGCGAGGCGCTATCGCTACCAAGTGCAGCCACGCA-----TGCTGTGTGCCGCG 1058
QY 694 TTTCTGTATGGAGGAGACGACATGTCAGGAGATTCAGGAGTTCACTCATGTGCGCG 753
Db 1059 TACCGGAGGGCAAGAGATGCCGTGTCAGGTGCTCAGGTGGTCCGCTGGTGTGCAAG 1118
QY 754 AATAAGAAAGGCGCTGCACTTGGCTGGTGTGACTTCTCTGGGGTTTGGGCTGTGGTGG 812
Db 1119 GCACTCAGTGGCGCTGGTTCTCTGCGGGGCTGTGCTCAGTGGGGCTGGGCTGTGGCG 1177

RESULT 14
US-09-798-403A-170
; Sequence 170, Application US/0978403A
; Publication No. US20030050240A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630P1C17
; CURRENT APPLICATION NUMBER: US/09/978.403A
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
```



APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J  
APPLICANT: Kijavini, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2630PIC25  
CURRENT APPLICATION NUMBER: US/09/978,564A  
CURRENT FILING DATE: 2001-10-16  
PRIOR FILING DATE: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366



**THIS PAGE BLANK (USPTO)**



```
Db 1751 GACGATCAACAGGCTCGCGAGCTTCTCTGGTGAGCAGGGATTTGGCTGGTTCGGCCGCC 1810
Qy 286 CACTGCATTGCAACACAGAAACATTGTGTCTACTTTTGAATGTTACTGTCTGGAG-----AG 339
Db 1811 CACTGCGTGTACGGGAGAAATATGAGCGCTTAAGTGAAGCAGTCTAGGCTGCAT 1870
Qy 340 TATGACTTAAGCCAGACAGACCCAGGAGAGCAAACTCTCACTATTGAACATGTTCATCATA 399
Db 1871 ATGGCATCAAAATCTGACTTCTCTCAGATAGAAACTAGGTTGATTGACCAAAATGTTCATA 1930
Qy 400 CATCCACATTTCTCCACCAAGAACCAATGACTATGATATGTCCTTTTGAAGATGCT 459
Db 1931 AACCCACA---CTACAATAACGGAAGAAACAATGACATTTGCCATGATCATCTTGAA 1987
Qy 460 GGAGCCTTCCAAATTTGGCCACTTTTGGGGCCCATATGCTTTCCAGAGCTCGGGGAGCAA 519
Db 1988 ATGAAGTGAACACTACACAGATTATATACAGCTTATTTGTTTACCAGAGAAATCAAGTT 2047
Qy 520 TTTGAGGCTGTTTTTATTTGTACAACTGACAGCTGGGGCCCTTTAACTGAAGGTGGCGTC 579
Db 2048 TTTCCCCCAGGAAGAAATTTGTTCTATTGCTGGCTGGGGGCACTTATATATCAAGGTTCT 2107
Qy 580 CTCCTCAAACTTTTCAGGAAGTGAATCTGCTTATTTGACCTGGGAAGAGTGTGGCA 639
Db 2108 ACTGCAGAGCTACTGCAAGAAGCTGACGTTCCCTTCTATCAATGAGAAATGCAACAA 2167
Qy 640 GCTCTGTTTAACTAAAGAGCCCATCAGTGGGAAGACCTTTCTTTGCACAGGTTTTTCT 699
Db 2168 CAGATGCCAGAAATAA-----CATTACGGAAATAATGTTGTGTCAGGCTATGAA 2218
Qy 700 GATGGAGGAGAGCGCATGTCAGGAGATTCAGGAGTTTCACTCATCTGTCGGGAATAAG 759
Db 2219 GCAGAGGGGTAGATCTTGTTCAGGGGATTCAGGCGGACCATCATGTGCGCAAGAAAC 2278
Qy 760 AAAGGGGCGCTGACTCTGGCTGGTGTGACTTCTCTGGGTTTGGGCTGTG 808
Db 2279 AACAG---ATGGCTCTGGCTGGCTGACGTCATTTGGATATCAATGTG 2324

RESULT 2
PCT-US94-00616-1
; Sequence 1, Application PC/TUS9400616
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: CLONING OF ENTEROKINASE AND METHOD OF USE
; NUMBER OF INVENTIONS: 33
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/00616
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2581 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
PCT-US94-00616-1
```

```
Query Match 7.7% Score 71; DB 5; Length 2581;
Best Local Similarity 49.08; Pred. No. 1.9e-13;
Matches 318; Conservative 0; Mismatches 310; Indels 21; Gaps 4;

Qy 166 ATCTTTGGAGGAGCAAGTGGAGAGGTTTCTTATCCCTGGCAGGTTATCTCTGAAACAA 225
Db 1691 ATTGTGGAGGAGTGAAGTACTCCAGAGAGGAGCCCTGGGCTTGGTCTGTCTGTTTC 1750
Qy 226 AGGCAGAGACATATTTTGGAGGAGCATGCTCTCACCACAGTGGGTGATCACCGGCGCT 285
```

```
Db 1751 GACGATCAACAGGCTCGCGAGCTTCTCTGGTGAGCAGGGATTTGGCTGGTTCGGCCGCC 1810
Qy 286 CACTGCATTGCAACACAGAAACATTGTGTCTACTTTTGAATGTTACTGTCTGGAG-----AG 339
Db 1811 CACTGCGTGTACGGGAGAAATATGAGCGCTTAAGTGAAGCAGTCTAGGCTGCAT 1870
Qy 340 TATGACTTAAGCCAGACAGACCCAGGAGAGCAAACTCTCACTATTGAACATGTTCATCATA 399
Db 1871 ATGGCATCAAAATCTGACTTCTCTCAGATAGAAACTAGGTTGATTGACCAAAATGTTCATA 1930
Qy 400 CATCCACATTTCTCCACCAAGAACCAATGACTATGATATGTCCTTTTGAAGATGCT 459
Db 1931 AACCCACA---CTACAATAACGGAAGAAACAATGACATTTGCCATGATCATCTTGAA 1987
Qy 460 GGAGCCTTCCAAATTTGGCCACTTTTGGGGCCCATATGCTTTCCAGAGCTCGGGGAGCAA 519
Db 1988 ATGAAGTGAACACTACACAGATTATATACAGCTTATTTGTTTACCAGAGAAATCAAGTT 2047
Qy 520 TTTGAGGCTGTTTTTATTTGTACAACTGACAGCTGGGGCCCTTTAACTGAAGGTGGCGTC 579
Db 2048 TTTCCCCCAGGAAGAAATTTGTTCTATTGCTGGCTGGGGGCACTTATATATCAAGGTTCT 2107
Qy 580 CTCCTCAAACTTTTCAGGAAGTGAATCTGCTTATTTGACCTGGGAAGAGTGTGGCA 639
Db 2108 ACTGCAGAGCTACTGCAAGAAGCTGACGTTCCCTTCTATCAATGAGAAATGCAACAA 2167
Qy 640 GCTCTGTTTAACTAAAGAGCCCATCAGTGGGAAGACCTTTCTTTGCACAGGTTTTTCT 699
Db 2168 CAGATGCCAGAAATAA-----CATTACGGAAATAATGTTGTGTCAGGCTATGAA 2218
Qy 700 GATGGAGGAGAGCGCATGTCAGGAGATTCAGGAGTTTCACTCATCTGTCGGGAATAAG 759
Db 2219 GCAGAGGGGTAGATCTTGTTCAGGGGATTCAGGCGGACCATCATGTGCGCAAGAAAC 2278
Qy 760 AAAGGGGCGCTGACTCTGGCTGGTGTGACTTCTCTGGGTTTGGGCTGTG 808
Db 2279 AACAG---ATGGCTCTGGCTGGCTGACGTCATTTGGATATCAATGTG 2324

RESULT 3
US-09-280-116-107
; Sequence 107, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
; FILE REFERENCE: 5800-24, 035800/176965
; CURRENT APPLICATION NUMBER: US/09/280,116A
; CURRENT FILING DATE: 1999-03-26
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 107
; LENGTH: 796
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: trypsin-like serine proteases
; NAME/KEY: misc_feature
; LOCATION: (1)..(796)
; OTHER INFORMATION: n = a, t, c or g
US-09-280-116-107

Query Match 5.9% Score 54; DB 4; Length 796;
Best Local Similarity 45.98; Pred. No. 4.7e-08;
Matches 220; Conservative 0; Mismatches 256; Indels 3; Gaps 1;

Qy 157 TTCAGTCCGATCTTTGGAGGAGCAAGTGGAGAGGTTTCTTATCCCTGGCAGGATATCT 216
Db 252 TCCAGCCGCAATGTTTGGTGGAGCTGTGCTCCCGAGGTTGAGTGGCCATGCGAGCCAGC 311
Qy 217 CTGAACAAGGCAAGACAGCATATTTTGGAGGAGCATGCTCTCACCACAGTGGGTGATC 276
```

Db 312 CTCAGGTCGGGTCGACACATCTGTGGGGGGCCCTCATCGCTGACCGCTGGGTGATA 371  
Qy 277 ACGGGGCTCACTGATTCGAAACAGAAACATTTGTCTACTTTTCAATTTACTCTCTGGA 336  
Db 372 ACAGTGGCCCACTGCTCCAGGAGGACAGATGGCCCTCCAGGGTCTGTGGACCGTGTTC 431  
Qy 337 GAGTATGACTTAAAGCCAGACAGACCCAGGAGAGCAAACTCT--CACTATTGAACACTGC 393  
Db 432 CTGGCAAGGTGTGCGACAACTCGCGCTGGCTGGAGAGGTGTCTTCAAGGTGAGCGGC 491  
Qy 394 ATCATACATCCACATTTCTCCACCAGAAACCAATGGACTATGATATTTGCCCTTTTGAAG 453  
Db 492 CTGCTCTCTGCAACCCGTACCAAGAGAGGACAGCCATGACTACGAGTGGCGTGTCTGAG 551  
Qy 454 ATGGCTGGAGCTTCCAAATTTGGCCACTTTGTGGGGCCCATATGCTTCTCCAGAGTGGG 513  
Db 552 CTCGACACCCGGTGGTGGCTGGCGCCGCTGGCGCCGCTGTGCTGCCCGCGCGCTCC 611  
Qy 514 GAGCAATTTGAGGCTGTTTTATTGTACAACTGAGGCTGGGCGCGCTTAACCTGAAGGT 573  
Db 612 CACTTCTCGAGCCGGCTGCACCTGCTGGATTAGGGCTGGGCGCTTGGCGAGGCG 671  
Qy 574 GCGTCTCTCACAGTCTTCAGAGAGTGAATCTGCCTATTTTGAACCTGGGAGAGTG 632  
Db 672 GCGCCCATCAGCAACGCTCTGCAGAACTGGATGTGCANTTGTATCCACAGACCTGTG 730

## RESULT 4

US-09-280-116-57  
; Sequence 57, Application US/09280116A  
; Patent No. 6331427

## GENERAL INFORMATION:

; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs  
; FILE REFERENCE: 5800-24, 035800/176965  
; CURRENT APPLICATION NUMBER: US/09/280.116A  
; CURRENT FILING DATE: 1999-03-26  
; NUMBER OF SEQ ID NOS: 268  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 57

LENGTH: 654

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: trypsin-like serine proteases

US-09-280-116-57

Query Match 5.7%; Score 52.6; DB 4; Length 654;  
Best Local Similarity 46.7%; Pred. No. 1.2e-07;  
Matches 203; Conservative 0; Mismatches 229; Indels 3; Gaps 1;

Qy 154 ATTTTCAGTCGCATCTTTGGAGGAAGCCAAAGTGGAGAGGGTCTCTATCCCTGGCAGGTA 213  
Db 118 ATGCTGAACCAATTTGGTGGGGGGGAGGACACGAGGAGGGCGAGTGGCCCTGGCAAGTC 177  
Qy 214 TCTCTGAAACAAAGCAGCAAGCATATTTGTGAGGAAGCATGCTCTACCACAGTGGGTG 273  
Db 178 AGCATCCAGCGCAACGGAAGCCACTTCTCGGGGGCAGCCCTCATCGCGGAGCAGTGGTGC 237  
Qy 274 ATCAGCGGGCTCACTGATTCGAAACAGAAACATTTGTCTACTTTGAATGTTACTGCT 333  
Db 238 CTGAGGGTGGGCACTGTCTCGCAACACCTCTGAGAGGTCTCCTGTACAGGTCTCTGTG 297  
Qy 334 GGAGAGTATGACTTAAGCCAGACAGACCCAGGAGAGCAAACTCTCACTATTGAACACTGC 393  
Db 298 GGGCAAGGCACTAGTGCACCGGGGACCAACAGCTATGTATGCCGGGTGAGGCAGGTG 357  
Qy 394 ATCATACATCCACATTTCTCCCAAGAAACCAATGGACTATGATATTTGCCCTTTTGAAG 453  
Db 358 GAGAGCAACCCCTGTACCAAGGACAGCGCTCCA--CGCTGAGCTGGCCCTGGTGGAG 414  
Qy 454 ATGGCTGAGCGCTTCCAAATTTGGCCACTTTGTGGGGGCCCATATGCTTCTCCAGAGTGGG 513

Db 415 CTGGAGGCAACGAGTGGCCCTTACCAATTAACATCTCCCGCTGCTGCTGACCCCTCG 474  
Qy 514 GAGCAATTTGAGGCTGGTTTTATTTGTACAACCTGCAGGCTGGGCGCTTAACTCAAGGT 573  
Db 475 GTGATCTTTGAGACGGGCATGAACCTGCTGGGTCACTGTGGGCGAGCCCCCAGTGAGGAA 534  
Qy 574 GCGCTCTCTCTCACAA 588  
Db 535 GACCTCTCTGCCGAA 549

## RESULT 5

US-09-386-653A-1  
; Sequence 1, Application US/09386653A  
; Patent No. 6458564

## GENERAL INFORMATION:

; APPLICANT: Andrade-Gordon, Patricia  
; APPLICANT: Darrow, Andrew  
; APPLICANT: Qi, Jian-shen  
; TITLE OF INVENTION: DNA encoding the novel human serine  
; FILE REFERENCE: ORT-1032  
; CURRENT APPLICATION NUMBER: US/09/386.653A  
; CURRENT FILING DATE: 1999-08-31  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1

LENGTH: 1110

TYPE: DNA

ORGANISM: Homo sapiens

US-09-386-653A-1

Query Match 5.7%; Score 52.6; DB 4; Length 1110;  
Best Local Similarity 46.7%; Pred. No. 1.7e-07;  
Matches 203; Conservative 0; Mismatches 229; Indels 3; Gaps 1;

Qy 154 ATTTTCAGTCGCATCTTTGGAGGAAGCCAAAGTGGAGAGGGTCTCTATCCCTGGCAGGTA 213  
Db 126 ATGCTGAACCAATTTGGTGGGGGGCAGGACACGAGGAGGGCGAGTGGCCCTGGCAAGTC 185  
Qy 214 TCTCTGAAACAAAGCAGCAAGCATATTTGTGAGGAAGCATGCTCTACCACAGTGGGTG 273  
Db 186 AGCATCCAGCGCAACGGAAGCCACTTCTCGGGGGCAGCCTCATCGCGGAGCAGTGGGTG 245  
Qy 274 ATCAGCGGGCTCACTGATTCGAAACAGAAACATTTGTCTACTTTGAATGTTACTGCT 333  
Db 246 CTGAGGGTGGGCACTGTCTTCGCAACACCTCTGAGAGCTCCCTGTACAGGTCTCTGTG 305  
Qy 334 GGAGAGTATGACTTAAGCCAGACAGACCCAGGAGAGCAAACTCTCACTATTGAACACTGTC 393  
Db 306 GGGCAAGGCACTAGTGCACCGGGGACCAACAGCTATGTATGCCGGGTGAGGCAGGTG 365  
Qy 394 ATCATACATCCACATTTCTCCCAAGAAACCAATGGACTATGATATTTGCCCTTTTGAAG 453  
Db 366 GAGAGCAACCCCTGTACCAAGGCGACGGCTCCA--CGCTGAGCTGGCCCTGGTGGAG 422  
Qy 454 ATGGCTGAGCGCTTCCAAATTTGGCCACTTTGTGGGGCCCATATGCTTCTCCAGAGTGGG 513  
Db 423 CTGAGGCAACGAGTGGCCCTTCAACCAATTAACATCTCTCCCGCTGTGCTGCTGACCCCTCG 482  
Qy 514 GAGCAATTTGAGGCTGGTTTTATTTGTACAACCTGCAGGCTGGGCGCTTAACTCAAGGT 573  
Db 483 GTGATCTTTGAGAGGGCATGAACCTGCTGGGTCACTGTGGGCGAGCCCCCAGTGAGGAA 542  
Qy 574 GCGCTCTCTCACAA 588  
Db 543 GACCTCTCTGCCGAA 557

## RESULT 6

US-09-386-653A-8  
; Sequence 8, Application US/09386653A







ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1) ..(1476)  
NAME/KEY: conflict  
LOCATION: (724)  
OTHER INFORMATION: Listed as T in GenBank Accession NO. U75329  
NAME/KEY: conflict  
LOCATION: (985)  
OTHER INFORMATION: Listed as C in GenBank Accession No. 6444419 U75329  
NAME/KEY: conflict  
LOCATION: (1347)  
OTHER INFORMATION: Listed as C in GenBank Accession No. 6444419 U75329  
NAME/KEY: conflict  
LOCATION: (1466)  
OTHER INFORMATION: Listed as A in GenBank Accession No. 6444419 U75329  
NAME/KEY: conflict  
LOCATION: (1471)  
OTHER INFORMATION: Listed as A in GenBank Accession No. 6444419 U75329.  
NAME/KEY: allele  
LOCATION: (478)  
OTHER INFORMATION: This base can be G or A with G being the more  
OTHER INFORMATION: common allele. The codon will change from Val to  
NAME/KEY: allele  
LOCATION: (777)  
OTHER INFORMATION: This base can be C or T with C being the more  
OTHER INFORMATION: common allele. The codon is unaffected with both  
OTHER INFORMATION: alleles encoding Gly.  
NAME/KEY: allele  
LOCATION: (768)  
OTHER INFORMATION: This base can be C or T with C being the more  
OTHER INFORMATION: common allele. This is a silent polymorphism.  
NAME/KEY: allele  
LOCATION: (834)  
OTHER INFORMATION: This base can be C or T with C being the more  
OTHER INFORMATION: common allele. This is a silent polymorphism.  
NAME/KEY: allele  
LOCATION: (625)  
OTHER INFORMATION: This base can be T or A with T being the more  
OTHER INFORMATION: common allele. The codon will change from Phe to Ile  
US-09-691-840-1

Query Match 5.6%; Score 51.4; DB 4; Length 1479;  
Best Local Similarity 49.0%; Pred. No. 5.1e-07;  
Matches 224; Conservative 0; Mismatches 226; Indels 7; Gaps 3;

QY 160 AGTCGATCTTTGGAGGAGCCAAAGTGGAGAGGGTTCCATCCCTGGCAGGTATCTCTG 219  
DB 760 AGCAGGATCGTGGCGCGAGAGCGCGCTCCCGGGGCGCTGGCCCTGGCAGGTACGCCGTG 819  
QY 220 AAACAAGGCAGAGCATATTTGTGAGGAAGCATGCTCTCACCAGTGGGTGATCAGC 279  
DB 820 CAGCTCCAGAGCTCCACGTGTGCGGAGGCTCCATCATCACCCTCCGAGTGGATCGTGACA 879  
QY 280 GGGCTCACTGCTATTCACAAACAAACATTTGTCTTA--CTTTGAATGTACTGCTGGAG 337  
DB 880 GCGGCCACTGCTGGAAAAACCTCTTAACAATCCATGCGATGGACGCATTTGGCGGG 939  
QY 338 AGTATGACTTAAGCCAGACAGACCCAGGAGCAAACTCTCACTAFTGAAACTGTCACTCA 397  
DB 940 ATTTTGAGACAATCTTTTCACTGTCTTATGAGCGCGGATACCAAGTA--GAAAGAAGTGATTT 997  
QY 398 TACATCCACATTTCTCACCAGAACCAATGAGTACTATGATATGTCCTTTTGAAGATGG 457  
DB 998 CTCATCAAAATTAATGACTCTCAAGA---CCAAAGAACATGACATTCGCTGATGAAGCTGC 1054  
QY 458 CTGGAGCCTTCCAAATTTGGCCACTTTGTGGGCCCATATGCTTCCAGAGCTGCGGAGC 517  
DB 1055 AGAAGCCTCTGACTTTCAACGACCTAGTGAACCAAGTGTCTGCGCCCAACCCAGGCATGA 1114  
QY 518 AATTTGAGGCTGCTTTTATTTGTACAACTGCAGGCTGGGCGCGCTTAACTGAAGGTGGCG 577

DB 1115 TGCTCAGCCAGAACAGCTCTCTGGATTTCGGGGTGGGGGCCACCGAGGAGAAAGGA 1174  
QY 578 TCCTCTCACAAGCTTGCAGGAAGTCAATCTGCCTAT 614  
DB 1175 AGACCTCAGAAGTGTGTAACGCTGCCAAGTGTCTT 1211

RESULT 12  
US-09-088-651-1  
; Sequence 1, Application US/09088651  
; Patent No. 6165771  
; GENERAL INFORMATION:  
; APPLICANT: BURGESS, NICOLA A.  
; APPLICANT: CLINKENBEARD, HELEN E.  
; APPLICANT: SOUTHAN, CHRISTOPHER D.  
; TITLE OF INVENTION: NOVEL COMPOUNDS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: RATNER & PRESTIA  
; STREET: P.O. BOX 980  
; CITY: VALLEY FORGE  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19482  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/088,651  
; FILING DATE: JUNE 1, 1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB9712088.5  
; FILING DATE: 10-JUNE-1997  
; APPLICATION NUMBER: EP 97308295.1  
; FILING DATE: 17-OCT-1997  
; APPLICATION NUMBER: GB 9803650.2  
; FILING DATE: 20-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PRESTIA, PAUL F.  
; REGISTRATION NUMBER: 23,031  
; REFERENCE/DOCKET NUMBER: GH30358  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-407-0700  
; TELEFAX: 610-407-0701  
; TELEX: 846169  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1109 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-09-088-651-1

Query Match 5.5%; Score 51; DB 4; Length 1109;  
Best Local Similarity 46.4%; Pred. No. 5.8e-07;  
Matches 202; Conservative 0; Mismatches 230; Indels 3; Gaps 1;

QY 154 ATTTTCAGTCGATTTCTTGGAGGAAGCCAAAGTGGAGAGGGTTCCCTATCCCTGGCAGGTA 213  
DB 287 ATGCTGAACCGAATGGTGGCGGGCAGGACACGAGGAGGCGGAGTGGCCCTGGCAAGTC 346  
QY 214 TCTCTGAAACAAAGGAGGAGGAGCATATTTGTGGAGGAAGCATGCTCTCACCACAGTGGGTG 273  
DB 347 AGCATCCACCGCAACGGAAGCCACTTCTCGGGGGCAGGCTCATCGGAGGAGTGGGTC 406  
QY 274 ATCAGCGGCGCTCACTGTCATTCGAAACAGAAATTTGTCTACTTTGAATGTTACTGCT 333  
DB 407 CTGACGGCTGCGCACTGCTTCCGCAACACCTCTGAGACGCTCCCTGGACGAGGTCTCTGCTG 466

QY	334	GGAGAGTATGACTTAAGCCACAGACACCAGAGAGCAAACTCTCACTATTGTAAACTGTC	393
DB	467	GGGCGAAGCAGTGTTAGTGACGCCGGGACACACCGTATGTATGC CGGGTGAGCGCAGGTG	526
QY	394	ATCATATATCCACATTTCTCCACCAGAAGAACCAATGGAGCTATGATATTGCCTTTTGAAG	453
DB	527	GAGAGCAACCCCTGTACACAGGCGACGGCCTCCAGCGCTTAGC---TGGCCCTTGGTGGAG	583
QY	454	ATGGCTGGAGCCCTTCCAATTTGGCCCACTTTGTGGGGGCCATATGCTCTTCAGAGCTGCGG	513
DB	584	CTGGAGCACCAAGTGCCTTTCACCAATTACATCTCCCGTGTGCTTGCCTGACCAAGGG	643
QY	514	GACAATTTGAGGCTGGTGTATTGTGTAACAAGCTGGGGGCCCTTAAGTGAAGGT	573
DB	644	CGAATTTTTGACAGGGCATGAAGTCTGGGTCACTGGCTGGGCGACCCCACTGAGGAA	703
QY	574	GCGTGCTCTCTCAAA	588
DB	704	GACTCTCTGCCGAA	718
RESULT 13			
US-09-386-629-2			
; Sequence 2, Application US/09386629			
; Patent No. 6426199			
; GENERAL INFORMATION:			
; APPLICANT: Darrow, Andrew L.			
; APPLICANT: Qi, Jenson			
; APPLICANT: Andrade-Gordon, Patricia			
; TITLE OF INVENTION: Identification and Characterization of the complementary			
; TITLE OF INVENTION: DNA encoding the novel human serine protease C-E			
; FILE REFERENCE: ORT-1030			
; CURRENT APPLICATION NUMBER: US/09/386,629			
; CURRENT FILING DATE: 1999-08-31			
; NUMBER OF SEQ ID NOS: 11			
; SOFTWARE: Patent In Ver. 2.0			
; SEQ ID NO 2			
; LENGTH: 1166			
; TYPE: DNA			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: Description of Artificial Sequence: C-E catalytic			
; OTHER INFORMATION: domain in a zymogen activated construct			
US-09-386-623-2			
Query Match 5.5%; Score 50.8; DB 4; Length 1166;			
Best Local Similarity 47.9%; Pred. No. 7e-07;			
Matches 315; Conservative 0; Mismatches 322; Indels 21; Gaps 5			
QY	166	ATTCTTTGGAGGAAGCCAACTGGAGAGGGTTCTCTATCCCTGGCAGGTATCTCTGAACAACRA	225
DB	166	ATCGTTTGGGGGTATGTCTCTAGAGGACAGCGAGTGGCCCTGGATCGTGACATCCAGAAG	225
QY	226	AGCAGAGAACATATTGTGGAGGAAGCATCGCTCCACACAGTGGGTGATCACGGCGGGCT	285
DB	226	AATGGGACCCACCACTGCGCAGGTCTCTGCTCACCAGCCGCTGGGTGATCATCTGTGCC	285
QY	286	CACGTG---CATTTGCAACAGAAAACATTTGTCTCTACTTTGAATGTTACTGCTGGAGAGTAT	342
DB	286	CACGTGTTTCAAGGACAACCTGAACAAAAACCATACCTGTCTCTGTGCTGCTGGGGGCTGG	345
QY	343	GACTTAAGCCAGACAGACCCAGGAGAGCAAACTCTCATAATTGNAACTGTCATCATACAT	402
DB	346	CAGCTGGGGAACCTTGGCTCTCGTTCGAGAGGTTGGGTGTTGCTGGGTGGAGCCCCAC	405
QY	403	CCACATTTTCTCCACCAAGAAAACCAATGGAGCTATGATATTGCCCTTTTGAAGATGGCTGGA	462
DB	406	CCTGTGTAATCTCTGGAAGGAGGTCCCTGTGCAACATTCGCCCTGGTGGCTCTCTGAGCGC	465
QY	463	GCCTTCCAATTTGGCCACTTTGTGGGGGCCCAATATGCTTTCAGAGCTGCGGGAGCAATTT	522
DB	466	TCCATACAGTTCCTCAGAGCGGGTCTCTGCCCATCTGGCTACTGATGCTCTATCCACCTC	525



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: March 25, 2003, 01:26:19 ; Search time 3423 Seconds  
(without alignments)  
2601.653 Million cell updates/sec  
Title: US-09-735-713a-2  
Perfect score: 1636  
Sequence: 1 MSLKMLISRNKLLILLGIVF.....IFTDISKVLSTWHEHIQTGN 306

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV=xlh  
-Q/cgn2\_1/USPTO\_spool/US09735713/runat\_18032003\_124141\_21590/app\_query.fasta\_1.455  
-DB-GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09735713 @CNG\_1.1.2496 @runat\_18032003\_124141\_21590 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg\_hum.\*  
31: em.htg\_inv.\*  
32: em.htg\_other.\*  
33: em.htg\_mus.\*  
34: em.htg\_pln.\*  
35: em.htg\_rtd.\*  
36: em.htg\_mam.\*  
37: em.htg\_vrt.\*  
38: em.sy.\*  
39: em.htgo\_hum.\*  
40: em.htgo\_mus.\*  
41: em.htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	1632	99.8	1262	6	AX342644	Sequence
2	1626	99.4	921	6	AX180293	Sequence
3	1626	99.4	1568	6	AX180299	Sequence
4	1608	98.3	909	6	AX180295	Sequence
5	1372.5	83.9	1671	6	AX360091	Sequence
6	898	54.9	495	6	AX180297	Sequence
7	831	50.8	164732	9	AC104237	Homo sapi
8	831	50.8	180707	2	AC012228	Homo sapi
9	688.5	42.1	3689	5	XL081291	Xenopus lae
10	657	40.2	3028	5	AB070367	Bufo japo
11	626	38.3	4628	5	XL081290	Xenopus lae
12	576	35.2	2627	6	AX480940	Sequence
13	529.5	32.4	169388	2	AC129620	Rattus no
14	496	30.3	942	3	SSU79521	Scolopendra
15	489.5	29.9	2409	6	AX360098	Sequence
16	486.5	29.7	2672	6	AX207905	Sequence
17	486.5	29.7	3104	6	AX207903	Sequence
18	486	29.7	2594	3	AY119618	Drosophila
19	486	29.7	3749	3	DROSTUBBLE	Drosophila
20	480.5	29.4	3183	10	BC029645	Mus muscu
21	476.5	29.1	708	6	AX207901	Sequence
22	475.5	29.1	1595	10	AB052292	Mus muscu
23	475.5	29.1	2231	10	BC019485	Mus muscu
24	469.5	28.7	2025	10	AF356627	Mus muscu
25	467	28.5	1359	6	AX354827	Sequence
26	467	28.5	1656	6	AX207907	Sequence
27	467	28.5	2412	6	AX375996	Sequence
28	467	28.5	2948	6	AX354825	Sequence
29	466	28.5	2267	6	AX335042	Sequence
30	466	28.5	2267	6	AX409602	Sequence
31	466	28.5	2267	9	HUMPPKKA	Nucleotide
32	464	28.4	2753	10	MUSPKA	Mouse plas
33	462.5	28.3	1801	3	AF357226	Panulirus
34	462	28.2	2752	10	MMU300738	Mus muscu
35	458	28.0	2398	10	BC026555	Mus muscu
36	456.5	27.9	1365	6	AX360088	Sequence
37	456.5	27.9	1365	6	AX392866	Sequence
38	456.5	27.9	2135	6	AX277411	Sequence
39	456.5	27.9	2135	9	AB038159	Homo sapi
40	456.5	27.9	2468	9	AB038157	Homo sapi
41	456.5	27.9	2554	9	AB038158	Homo sapi
42	456.5	27.9	2951	6	AX392864	Sequence
43	454.5	27.8	3387	6	AX360082	Sequence
44	450	27.5	1743	10	BC003851	Mus muscu
45	450	27.5	1753	10	AF378085	Mus muscu

ALIGNMENTS

RESULT 1

AX342644 AX342644 1262 bp DNA linear PAT 12-JAN-2002  
LOCUS Sequence 41 from Patent WO0198468.  
ACCESSION AX342644  
VERSION AX342644.1 GI:18152041  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
AUTHORS Yue,H., Elliott,V.S., Gandhi,A.R., Lal,P., Au-Young,J.,  
Tribouley,C.M., Deleage,A.M., Baughn,M.R., Nguyen,D.B., Lee,E.A.,  
Hafalia,A., Khan,F.A., Walia,N.K., Yao,M.G., Lu,D.A., Patterson,C.,  
Tang,Y.T., Walsh,R.T., Azimzai,Y., Ramkumar,J., Xu,Y. and Reddy,R.  
JOURNAL Patent: WO 0198468-A 41 27-DEC-2001;  
Incyte Genomics, Inc. (US)  
FEATURES  
source  
1. 1262  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 433459CB1"  
BASE COUNT 354 a 265 c 306 g 337 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.15e-144 Length: 1262  
Score: 1632.00 Matches: 305  
Percent Similarity: 99.67% Conservative: 0  
Best Local Similarity: 99.67% Mismatches: 1  
Query Match: 99.76% Indels: 0  
DB: 6 Gaps: 0  
US-09-735-713A-2 (1-306) x AX342644 (1-1262)  
Qy 1 MetSerLeuLysMetLeuLysSerArgAsnLysLeuLysLeuLeuLeuGlyIleValPhe 20  
Db 165 ATGAGTCTCAAAATGCTTATAGCAGGAACAAGCTGATTACTACTAGGAATAGCTTT 224  
Qy 21 PheGluArgGlyLysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyClnSer 40  
Db 225 TTGACAGAGGTAAATCTGCAACTCTTTCGCTCCCAAGCTCCAGTGTGGGAGAGT 284  
Qy 41 LeuValLysValGlnProTrpAsnTyrPheAsnIlePheSerArgIleLeuGlyGlySer 60  
Db 285 CTGGTTAAGGTACAGCCCTTGAATATTATTAACATTTTCAGTCGCATTTCTGGAGGAGC 344  
Qy 61 GlnValGluLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80  
Db 345 CAAGTGGAAGAAGGTTCTATCCCTGGCAGGTATCTCTGAACAACAAAGGAGAGCATATT 404  
Qy 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100  
Db 405 TGTGGAGAGCATGCTCTACACAGTGGGTGATCAGCGGGCTCACTGCATTGCAAAAC 464  
Qy 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120  
Db 465 AGAAACATTTGCTACTTTCAATGTTTACTGCTGAGAGGTATGACTTAAAGCCAGACAGAC 524  
Qy 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140  
Db 525 CCAGAGAGCAAACTCTCACATTATGAAACTGTCATPACATCCACATTTCTCCACCAG 584  
Qy 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160  
Db 585 AAACCAATGGACTATGATATGCCCTTTTGAAGTGGCTGGAGCTTCCAATTTGGCCAC 644  
Qy 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180  
Db 645 TTTGTGGGGCCCATATGCTCTCCAGAGCTGGGGAGCAATTTGAGGCTGTTTATTGT 704  
Qy 181 ThrThrAlaGlyTyrGlyArgLeuThrClnGlyGlyValLeuSerGlnValLeuGlnCln 200  
|||||

Db 705 ACAACTGCAGGCTGGGGCCGCTTAACCTGAAGGTGGCTCTCTCAACAAGCTTTGCAGGAA 764  
Qy 201 ValAsnLeuProIleLeuThrTrpGluCysValAlaAlaLeuLeuThrLeuLysArg 220  
|||||  
Db 765 GTGAATCTGCCTATTTTACCTGGGAAGAGTGTGGGAGCTCTGTTAACACTAAAGAGG 824  
Qy 221 ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys 240  
Db 825 CCCATCAGTGGGAAGACCTTTCTTTCACAGGTTTCTGATGGAGGAGAGACGCATGT 884  
Qy 241 GlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAla 260  
|||||  
Db 885 CAGGAGAGATTCAGGAGGTCTACTCATGTGCGGAATAAGAAAGGGCCCTGGACTCTGGCT 944  
Qy 261 GlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTyrPheAsnValArgLysSer 280  
|||||  
Db 945 GGTGTGACTTCTGGGGTTTGGGTGTGCTGAGGCTGGAGAAACAATGTGAGGAAAGT 1004  
Qy 281 AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu 300  
Db 1005 GATCAAGGATCCCTGGGATCTTCACAGACATAGTAAGTGTCTTCTGGATCCACGAA 1064  
Qy 301 HisIleGlnThrGlyAsn 306  
Db 1065 CACATCCAAACTGGTAAC 1082  
RESULT 2  
AX180293  
LOCUS AX180293 921 bp DNA linear PAT 06-AUG-2001  
DEFINITION Sequence 1 from Patent WO0146407.  
ACCESSION AX180293  
VERSION AX180293.1 GI:15132262  
KEYWORDS  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 921)  
AUTHORS Walke,D.W., Turner,C.A., Abuin,A., Friedrich,G., Zambrowicz,B. and  
Sands,A.T.  
TITLE Polynucleotides encoding human,protease homologs  
JOURNAL Patent: WO 0146407-A 1 28-JUN-2001;  
Lexicon Genetics Incorporated (US)  
FEATURES  
Location/Qualifiers  
source  
1. 921  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
BASE COUNT 242 a 195 c 240 g 242 t 2 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.9e-144 Length: 921  
Score: 1626.00 Matches: 304  
Percent Similarity: 99.35% Conservative: 0  
Best Local Similarity: 99.35% Mismatches: 2  
Query Match: 99.39% Indels: 0  
DB: 6 Gaps: 0  
US-09-735-713A-2 (1-306) x AX180293 (1-921)  
Qy 1 MetSerLeuLysMetLeuLysSerArgAsnLysLeuLeuLeuGlyIleValPhe 20  
Db 1 ATGAGTCTCAAAATGCTTATAGCAGGAACAAGCTGATTACTACTAGGAATAGCTTT 60  
Qy 21 PheGluArgGlyLysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyClnSer 40  
Db 61 TTTGAACRAGGTAAATCTGCACTCTTTCGCTCCCAAGCTCCAGTGTGGGAGAGT 120  
Qy 41 LeuValLysValGlnProTrpAsnTyrPheAsnIlePheSerArgIleLeuGlyGlySer 60  
Db 121 CTGGTTAAGGTACAGCTTGAATATTATTAACATTTTCAGTCGCATTTCTGGAGGAGC 180  
Qy 61 GlnValGluLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80  
|||||

Db	181	CAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAAGGAGAGCATATT	240
QY	81	CysGlyGlySerThrLeuValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn	100
Db	241	TGTGGAGGAGCATCGCTCACACAGTGGGTGATGATCGCGGCTCCTGCTATGCAAC	300
QY	101	ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp	120
Db	301	AGAAACATTGTGTCTACTTTGAATGTTACTGTGGAGAGATGATGACTTAAGCCAGACAGAC	360
QY	121	ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys	140
Db	361	CCAGGAGAGCAACTCTCACTATTGAACCTGTCATCATCATCATCTTCCACCAAG	420
QY	141	LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis	160
Db	421	AAACCAATGGACATGATGATATGCGCTTTTGAAGATGGCTGGAGCCCTTCAATTTGCCAC	480
QY	161	PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys	180
Db	481	TTTGTGGGCCCATATGCTTCCAGAGCTGCGGAGCAATTTGAGGCTGGTTTATTGT	540
QY	181	ThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnGlu	200
Db	541	ACAACCTGCAGGCTGGGGCCGCTTAACCTGAAGGTGGCGTCTCTCACAAAGTCTTGCAGGA	600
QY	201	ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg	220
Db	601	GTGAATCTGCCATATTTTGACCTGGGAGAGTGTGTGGCAGCTCTGTAAACACTAAAGAGG	660
QY	221	ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys	240
Db	661	CCCATCAGTGGGAAGACCTTCTTTCACAGAGTTTTCCTCATGGAGGAGAGACGATGT	720
QY	241	GlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysGlyAlaTrpThrLeuAla	260
Db	721	CAGGAGATTCAGGAGGTTCACTCATGTGCCGAATAAGAAAGGCGCTGGACTCTGGCT	780
QY	261	GlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSer	280
Db	781	GGTGTCACTCTCCGTGGGTTTGGCTGTGTGGAGGCTGGAGAAACAATGTGAGAGAAAGT	840
QY	281	AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu	300
Db	841	GATCAAGGATCCCTGGGATCTTCACAGACATTAGTAAAGTGCTTTCCTGGATCCAGAA	900
QY	301	HisIleGlnThrGlyAsn	306
Db	901	CACATCCAAACTGGTAAC	918
RESULT 3			
AX180299			
LOCUS	AX180299	1568 bp	DNA linear PAT 06-AUG-2001
DEFINITION	Sequence 7 from Patent WO0146407.		
ACCESSION	AX180299		
VERSION	AX180299.1 GI:15132265		
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1. (bases 1 to 1568)		
AUTHORS	Walke,D.W., Turner,C.A., Abuin,A., Friedrich,G., Zambrowicz,B. and Sands,A.I.		
TITLE	Polynucleotides encoding human protease homologs		
JOURNAL	Patent: WO 0146407-A 7 28-JUN-2001;		
FEATURES	Lexicon Genetics Incorporated (US)		
source	Location/Qualifiers		
	1..1568		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
BASE COUNT	452 a	338 c	353 g 423 t 2 others

ORIGIN			
Alignment Scores:			
Pred. No.:	5,47e-144	Length:	1568
Score:	1626.00	Matches:	304
Percent Similarity:	99.35%	Conservative:	0
Best Local Similarity:	99.35%	Mismatches:	2
Query Match:	99.39%	Indels:	0
DB:	6	Gaps:	0
US-09-735-713A-2 (1-306) x AX180299 (1-1568)			
QY	1	MetSerLeuLysMetLeuIleSerArgAsnLysLeuIleLeuLeuGlyIleValPhe	20
-Db	245	ATGAGTCTCAAAATGCTTATAAGCAGGAACAAGCTGATTTTACTACTAGGAATAGTCTTT	304
QY	21	PheGluArgGlyLysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGlnSer	40
Db	305	TTTGAACRAGGTAAATCTGTCARCTCTTTCGCTCCCAAGCTCCAGTTGTGGCAGAGT	364
QY	41	LeuValLysValGlnProTrpAsnTyrPheAsnIlePheSerArgIleLeuGlyGlySer	60
Db	365	CTGTTAAGGTACAGCTTGGAAATATTTAACATTTTTCAGTCGATTTCTTGGAGGAGC	424
QY	61	GlnValGluLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle	80
Db	425	CAAGTGGAGAGGGTTCCTATCCCTGCGCAGGTATCTCTGAACAAAGGCAAGACATATT	484
QY	81	CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn	100
Db	485	TGTGGAGGAGACATCGTCTCACACAGTGGGTGATCAGCGCGCTCCTCATTTGCAAAC	544
QY	101	ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp	120
Db	545	AGAAACATTGTGTCTACTTTGAATGTTACTGCTGGAGAGTATGACTTAACCCAGACAG	604
QY	121	ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys	140
Db	605	CCAGGAGAGCAAACTCTCACTATTGAAACTGTCTATCATCATCATCATTTTCCACCAAG	664
QY	141	LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis	160
Db	665	AAACCAATGGACATGATGATTTGCCCTTTTGAAGATGGCTGGAGCTTCCAAATTTGGCCAC	724
QY	161	PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys	180
Db	725	TTTGTGGGCCCATATGCTTCCAGAGCTGCGGAGGCAATTTGAGGCTGGTTTATTGT	784
QY	181	ThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnGlu	200
Db	785	ACAACTGCAGGCTGGGGCCGCTTAACTGAAGTGGCGTCTCTCACAAAGTCTTGCAGGAA	844
QY	201	ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg	220
Db	845	GTGAATCTGCTTATTTTGACCTGGGAGAGTGTGTGGCAGCTCTGTTTAAACACTAAAGAGG	904
QY	221	ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys	240
Db	905	CCCATCAGTGGAGAGACCTTCTTTCGACAGGTTTTCCTGATGGAGGAGAGACGATGT	964
QY	241	GlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAla	260
Db	965	CAGGAGATTCAGGAGGTTTCACTCATGTCGCGGAATAAGAAAGGGGCTGAGCTCTGGCT	1024
QY	261	GlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSer	280
Db	1025	GGTGTCACTTCTCCGTGGGCTTGGGCTGTGTGCGAGGCTGGAGAAACAATGTGAGGAAAGT	1084
QY	281	AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu	300
Db	1085	GATCAAGGATCCCTGGGATCTTTCACAGACATTAGTAAAGTGCTTTCCTGGATCCAGAA	1144
QY	301	HisIleGlnThrGlyAsn	306

```
Db 1145 CACATCAAACTGGTAAC 1162
|||||
RESULT 4
AX180295
LOCUS AX180295 909 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 3 from Patent WO0146407.
ACCESSION AX180295
VERSION AX180295.1 GI:15132263
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 909)
AUTHORS Waikie D.W., Turner, C.A., Abuin, A., Friedrich, G., Zambrowicz, B. and
Sands, A.T.
TITLE Polynucleotides encoding human protease homologs
JOURNAL Patent: WO 0146407-A 3 28-JUN-2001;
Lexicon Genetics Incorporated (US)
FEATURES
source
Location/Qualifiers
1..909
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 237 a 193 c 238 g 239 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 1.42e-142 Length: 909
Score: 1608.00 Matches: 300
Percent Similarity: 99.34% Conservative: 0
Best Local Similarity: 99.34% Mismatches: 2
Query Match: 98.29% Indels: 0
DB: 6 Gaps: 0
US-09-735-713a-2 (1-306) x AX180295 (1-909)
QY 5 MetLeuIleSerArgAsnLysLeuIleLeuLeuGlyIleValPheGluArgGly 24
|||||
Db 1 ATGCTTATAAGCAGGAACAGGTGATTTTACTACTAGGAATAGTCTTTTGAACRAGGT 60
QY 25 LysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGlnSerLeuValysVal 44
|||||
Db 61 AAATCTGCARCTCTTCGCTCCCAAGATCCCGAGTGTGGCAGAGTCTGGTTAAGGTA 120
QY 45 GlnProTyrAsnTyrPheAsnIlePheSerArgIleLeuGlyGlySerGlnValGluLys 64
|||||
Db 121 CAGCCTTGAATATTTTAAACATTTTCAGTCGATCTTGGAGGAAGCAAGTGGAGAG 180
QY 65 GlySerTyrProTyrGlnValSerLeuLysGlnArgGlnLysHisIleCysGlyGlySer 84
|||||
Db 181 GGTTCCTATCCCTGGCAGGTATCTGAAACAAAGCAGAGCATATTTGTGGAGGAAGC 240
QY 85 IleValSerProGlnTyrValIleThrAlaAlaHisCysIleAlaAsnArgAsnIleVal 104
|||||
Db 241 ATCGTCTCACACAGTGGTGATCAGCGGGCTACTCGATTCGAAACAGAAACATTTGTG 300
QY 105 SerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAspProGlyGluGln 124
|||||
Db 301 TCTACTTTGAATGTTACTGCTGGAGAGTATGACTTAAGCCAGACAGACCCAGGAGGAA 360
QY 125 ThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLysLysProMetAsp 144
|||||
Db 361 ACTCTCACTATTTGAACCTGTATCATCATCATCATCATTTCTCCACCAAGAAACCAATGGAC 420
QY 145 TyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHisPheValGlyPro 164
|||||
Db 421 TATGATATGGCCCTTTTGAAGATGGCTGGAGCCCTTCCAAATTTGGCCACTTTGTGGGGCCC 480
QY 165 IleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThrAlaGly 184
|||||
Db 481 ATATGCTCTCCAGAGCTCGGGAGCAATTTGAGGCTGGTTTATTTTGTACAACTGCAGGC 540
|||||
```

```
QY 185 TrpGlyArgLeuThrGluGlyValLeuSerGlnValLeuGlnGluValAsnLeuPro 204
|||||
Db 541 TGGGCGCGCTTAACCTGAAGGTGGCGCTCTCTCACAGTCTTCAGAGAGTGAATCTGCCT 600
QY 205 IleLeuThrTrpGluGluCysValAlaLeuLeuThrLeuLysArgProIleSerGly 224
|||||
Db 601 ATTTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTTAAACACTAAAGAGGCCATCAGTGGG 660
QY 225 LysThrPheLeuCysThrClyPheProAspGlyGlyArgAspAlaCysGlnGlyAspSer 244
|||||
Db 661 AAGACCTTTCTTTGTCACAGGTTTCTGATGAGGAGAGAGCATGTCTCAGGAGATTCA 720
QY 245 GlyGlySerLeuMetCysArgAsnLysGlyAlaTrpThrLeuAlaGlyValThrSer 264
|||||
Db 721 GGAGGTTCACTCATGTGCGGAATAAGAAAGGGGCTGGACTCTGGCTGGTGTGACTTCC 780
QY 265 TrpGlyLeuGlyCysGlyArgGlyTrpArgAsnAsnValArgLysSerAspGlnGlySer 284
|||||
Db 781 TGGGGTTTGGGCTGTGGTCGAGGCTGGAGAAACAATGTGAGGAAAGTGTCAAGGATCC 840
QY 285 ProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThr 304
|||||
Db 841 CCTGGGATCTTCACAGACATTAAGTAAAGTGCTTCTCTGGATCCACAGACATCAAACT 900
QY 305 GlyAsn 306
|||||
Db 901 GGTAAC 906
RESULT 5
AX360091
LOCUS AX360091 1671 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 47 from Patent WO0200860.
ACCESSION AX360091
VERSION AX360091.1 GI:18675717
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Plowman, G., Whyte, D., Sudarsanam, S., Manning, G., Caenepeel, S. and
Charyczak, G.
TITLE Novel proteases
JOURNAL Patent: WO 0200860-A 47 03-JAN-2002;
Sugen, Inc. (US)
FEATURES
source
Location/Qualifiers
1..1671
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 451 a 369 c 404 g 447 t
ORIGIN
Alignment Scores:
Pred. No.: 4.93e-120 Length: 1671
Score: 1372.50 Matches: 266
Percent Similarity: 93.75% Conservative: 4
Best Local Similarity: 92.36% Mismatches: 6
Query Match: 83.89% Indels: 12
DB: 6 Gaps: 2
US-09-735-713a-2 (1-306) x AX360091 (1-1671)
QY 1 MetSerLeuLysMetLeuIleSerArgAsnLysLeuIleLeuLeuGlyIleValPhe 20
|||||
Db 1 ATGAGTCTCAAAATGCTTTAAGCAGGAGCAACAGCTGATTTTACTACTAGGAATAGTCTTT 60
QY 21 PheGluArgGlyLysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGlnSer 40
|||||
Db 61 TTTGAACGAGGTAAATCTGCAACTCTTTTCGCTCCCAAGACTCCCAAGTGTGTGGCAGAGT 120
QY 41 LeuValLysValGlnProTyrAsnTyrPheAsnIlePheSerArgIleLeuGlyGlySer 60
|||||
Db 121 CTGGTTAAGGTACAGCCTTGGAAATATTTTAACTATTTTTCAGTCGCATTTCTTGGAGGAAGC 180
```

```
QY 61 GlnValGluLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80
|||||
Db 181 CAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAACAAGGAGGAGCATATT 240
|||||
QY 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaHisCysIleAlaAsn 100
|||||
Db 241 TGTGGAGGAGCATCGTCTCACCACAGTGGGTGATCAGCGCGCTCAGTCATTTGCAAC 300
|||||
QY 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120
|||||
Db 301 AGAAACATTTGTCTACTTTGAATGTACTCTGGAGAGTATCAGTTAAGCCAGACAGAC 360
|||||
QY 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140
|||||
Db 361 CCAGGAGAGCAACTCTCACTATTGAACGTCTCATCATCATCATCTCTCCACCAAG 420
|||||
QY 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160
|||||
Db 421 AAACCAATGGACTATGATATTGCCCTTTTGAAGATGGCTGGAGCCTTCCCAATTTGGCCAC 480
|||||
QY 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180
|||||
Db 481 TTTGTGGGGCCCATATGCTCTCCAGAGCTGCGGGAGCAATTTTCAGGCTGGTTTTATTGT 540
|||||
QY 181 ThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnGlu 200
|||||
Db 541 ACAACTGCAGCTGGGGCCCTTAACCTGAAGGTGGCGTCTCTCAACAAGTCTTTCAGGAA 600
|||||
QY 201 ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg 220
|||||
Db 601 GTCAATCTGCCTATTTTGACCTGGGAAGAGTGTGTGCAGCTCTGTTAACTAAAGAGG 660
|||||
QY 221 ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys 240
|||||
Db 661 CCCATCAGTGGGAGAGCTTCTTTTCACAGGTTTTCCTGATGGAGGAGAGCAGCATGT 720
|||||
QY 241 GlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrp-ThrLeuAl 260
|||||
Db 721 CAGGGAGATTCAGAGGTTCCATCTGTCGCGGAATAAGAAAGGGCCCTGGGATCTGGC 780
|||||
QY 260 aGlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSe 280
|||||
Db 781 TGGT-----CAATTTGGG-----AGGCTCAGGTGGGA 807
|||||
QY 280 rAspGlnGlySerProGlyIle 287
|||||
Db 808 GGATCGCTTGAGTCCAGGAGTT 829
|||||
RESULT 6
LOCUS AX180297 495 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 5 from Patent WO0146407.
ACCESSION AX180297
VERSION AX180297.1 GI:15132264
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Walke,D.W., Turner,C.A., Abuin,A., Friedrich,G., Zambrowicz,B. and
Sands,A.T.
TITLE Polynucleotides encoding human protease homologs
JOURNAL Patent: WO 0146407-A 5 28-JUN-2001;
Lexicon Genetics Incorporated (US)
FEATURES
Source 1. 495
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 117 a 101 c 147 g 130 t
ORIGIN
```

```
Alignment Scores: 6.87e-76 Length: 495
Score: 898.00 Matches: 164
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 54.89% Indels: 0
DB: Gaps: 0
US-09-735-713A-2 (1-306) x AX180297 (1-495)
QY 143 MetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHisPheVal 162
|||||
Db 1 ATGGACTATGATATTGCCCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGCCACTTTGTG 60
|||||
QY 163 GlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThr 182
|||||
Db 61 GGCCCATATGCTTCCAGAGCTGCGGGAGCAATTTTGGAGCTGGTTTTATTGTGTACAACT 120
|||||
QY 183 AlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnValAsn 202
|||||
Db 121 GCAGGCTGGGGCCGCTTAACCTGAAGGTGGCGTCTCTCAACAAGTCTTGCAGGAAGTGAAT 180
|||||
QY 203 LeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArgProIle 222
|||||
Db 181 CTGCCATTTTGGACCTGGGAAGAGTGTGTGCAGCTCTGTTAACTATAAGAGGCCCATC 240
|||||
QY 223 SerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgaspAlaCysGlnGly 242
|||||
Db 241 AGTGGGAAGACCTTCTTTCACAGGTTTTTCTGATGGAGGAGAGACGATGTCAGGGA 300
|||||
QY 243 AspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAlaGlyVal 262
|||||
Db 301 GATTCAGGAGGTTCACTCATGTCCGGAATAAGAAAGGGCCCTGGGACTCTGGCTGGTGTG 360
|||||
QY 263 ThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSerAspGln 282
|||||
Db 361 ACTTCCTGGGTTTGGGCTGTGTCGAGGCTGGAGAAACAATGTGAGAGAAAGTGAATCAA 420
|||||
QY 283 GlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIle 302
|||||
Db 421 GSATCCCTGGGATCTTTCACAGACATTAGTAAAGTGCTTCTTCTGGATCCAGAACATC 480
|||||
QY 303 GlnThrGlyAsn 306
|||||
Db 481 CAACCTGGTAAAC 492
|||||
RESULT 7
LOCUS AC104237 164732 bp DNA linear PRI 30-APR-2002
DEFINITION Homo sapiens chromosome 11, clone RP11-35J10, complete sequence.
ACCESSION AC104237
VERSION AC104237.2 GI:20128277
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 11, clone RP11-35J10
JOURNAL Unpublished
REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,W., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
```





Db 51007 GGCCAGACTTCTCTGAGTCTGCAGCATGAGAGAGAGAGAAAGAGAAAGTCTGCTGCTTT 51066  
 QY 240 ----- 240  
 Db 51067 GATGAAGCCATGATGAGAGAAAGGGAGGTGGGATTGTTGGTTCACAGTGATAGTTGAG 51126  
 QY 240 ----- 240  
 Db 51127 GATCATATGGGAATTGCTGCTGGATCTGTATACAGGACCCTGAGTCTAAATTATTAGA 51186  
 QY 240 ----- 240  
 Db 51187 ACTGAGGAGATCATGATCTCTCCAGCATCCTTGTGAGGAACACAGTGAGCATTTTG 51246  
 QY 241 ----- GlnGlyAspSerG1 245  
 Db 51247 GGGCATAGACACTGACAGTGTATAGAGGCTTTTCCCTCTGCAGGGAGATTCAGG 51306  
 QY 245 YGlySerLeuMetCysArgAspLysGlyAlaTrpHrLeuAlaGlyValThrSerTr 265  
 Db 51307 AGGTTCACTCATGTGCCGGAATAAGAAAGGGCCCTGGACTCTGGCTGTGACTTCCTG 51366  
 QY 265 pGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSerAspGlnGlySerPr 285  
 Db 51367 GGGTTGGCTGTGCTGAGGCTGAGAAACAATCTGAGGAAAGTGTATCAGGATCCCC 51426  
 QY 285 oGlyllePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrG1 305  
 Db 51427 TGGGATCTTCACAGACATAGTAAAGTGTCTCCCTGGATCCACGAACACATCCAAACTGG 51486  
 QY 305 yAsn 306  
 Db 51487 TAAC 51490

RESULT 8  
 AC012228/c  
 LOCUS Homo sapiens chromosome 11 clone RP11-439A13 map 11, \*\*\* SEQUENCING  
 IN PROGRESS \*\*\*, 2 ordered pieces.  
 AC012228  
 AC012228.8 GI:22123492  
 HTG: HTGS\_PHASE2; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 AUTHORS Birren, B., Nusbaum, C. and Lander, E.  
 TITLE Homo sapiens chromosome 11, clone RP11-439A13  
 JOURNAL Unpublished  
 AUTHORS 2 (bases 1 to 180707)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,  
 Baldwin, J., Barna, N., Beckerly, R., Boguslavskiy, L., Boukhgalter, B.,  
 Brown, A., Castle, A., Colangelo, M., Collins, S., Collumore, A.,  
 Cooke, P., Dearellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,  
 Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D.,  
 Galagan, J., Gardyna, S., Grant, G., Hagsos, B., Heaford, A., Horton, L.,  
 Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,  
 Lehoczy, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,  
 McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,  
 Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,  
 Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,  
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
 Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,  
 Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

Direct Submission  
 Submitted (21-OCT-1999) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 180707)  
 Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,  
 Barna, N., Bastien, J., Bloom, T., Boguslavskiy, L., Boukhgalter, B.,  
 Canarata, J., Chang, J., Chazaro, B., Choepel, Y., Collumore, A.,  
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,

Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,  
 Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagsos, B.,  
 Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,  
 Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K.,  
 Liu, G., MacLean, C., Macdonald, P., Major, J., Matthews, C.,  
 McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,  
 Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H.,  
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
 Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
 Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,  
 Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,  
 Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,  
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,  
 Zembek, L., Zimmer, A. and Zody, M.

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WBIR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L2182  
 Center clone name: 439\_A\_13

NOTE: This is a 'working draft' sequence. It currently  
 consists of 2 contigs. Gaps between the contigs  
 are represented as runs of N. The order of the pieces  
 is believed to be correct as given, however the sizes  
 of the gaps between them are based on estimates that have  
 been provided by the submitter.  
 This sequence will be replaced  
 by the finished sequence as soon as it is available and  
 the accession number will be preserved.  
 1 71362: contig of 71362 bp in length  
 71363 71462: gap of 100 bp  
 71463 180707: contig of 109245 bp in length.

Location/Qualifiers  
 1. 180707  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="11"  
 /map="11"  
 /clone="RP11-439A13"  
 /clone\_lib="RPC1-11 Human Male BAC"  
 47741 a 39327 c 40651 g 52888 t 100 others

BASE COUNT  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.67e-66 Length: 180707  
 Score: 831.00 Matches: 205  
 Percent Similarity: 30.21% Conservative: 1  
 Best Local Similarity: 30.06% Mismatches: 1  
 Query Match: 50.79% Indels: 475  
 DB: 2 Gaps: 3

US-09-735-713A-2 (1-306) x AC012228 (1-180707)

QY 100 AsnArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyAspLeuSerGlnThr 119  
 Db 167186 CATGAAACATGTGCTACTTTGAATGTTACTGCTGGAGAGATGACTTAAAGCAGACA 167127  
 QY 120 AspProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThr 139  
 Db 167126 GACCCAGGAGAGCAAACTCTCACATTGCAAACTGTCATCATCATCATCTTCCACC 167067  
 QY 140 LysLysProMetAspTyAspIleAlaLeuLysMetAlaGlyAlaPheGlnPhe--- 158  
 |||||

TITLE  
JOURNAL

COMMENT

Db 167066	AGAAACCAATGGACTATGATATTGCCCTTTGAAGATGCTGGAGCCTTCCAATTTGCT	167007
QY 158	-----	158
Db 167006	AAACATTTGAGGATCATACTGAATGTCTACCCCATGTTCTCTATAACCTCTGTGCATA	166947
QY 158	-----	158
Db 166946	GCACATTTCTCGAGTATACGTTCTGGGTCTGCACATTTGAGATAGGAATTTAGGCTGGA	166887
QY 158	-----	158
Db 166886	GAGTCTACCTGAAACTCTTTTGTGTTCTGGGTCTGGAGATAGCATTACATGGAGTCAGG	166827
QY 158	-----	158
Db 166826	AGAAGGTTTCTTAATTGACAAGTAACCCCTGGAGGTGCCAGGCAGAGCTATCACAACT	166767
QY 158	-----	158
Db 166766	CTCATGAGTCTTTAGGGGACCACCTGGATACCAAGTGTGAGTGGCTTATAGTGTCTC	166707
QY 159	-----GlyHisPheValGlyPro 164	
Db 166706	TCCTCCATGAATGACACAAGCCCTACCTTGCCTTTTCTTTCAGGCCACTTTGTGGGGCC	166647
QY 165	IleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThrAlaGly 184	
Db 166646	ATATGTTCTCCAGAGCTCGGGAGCAATTTGAGGCTGGTTTTATTGTACAACATGCAGGC	166587
QY 185	TrpGlyArgLeuThrGlu-----	190
Db 166586	TGGGGCCGCTTAACGGAAGTAAGRACTTGTGTGCCATTCCCTTAAGCAGAGGTCAAGAA	166527
QY 190	-----	190
Db 166526	CTTCTGTGGGCAATCAGAAAATGGAATTTTAAATCCATTTTAAAAATCGATTGTGCA	166467
QY 190	-----	190
Db 166466	TACATTTCTAAGATAGGAANTGCATGTTACTTTAGAAAATCAAGTCTTGCCAAAGCGA	166407
QY 190	-----	190
Db 166406	CACAAACAAATGATGCATCTTCAACTCAGAAAAGGTGAAAGAGTTCTGCTGGGGCTT	166347
QY 191	-----GlyGlyValLeuSerGlnValLeuGlnValLeuGlnValLeuPro 204	
Db 166346	CCATCTCTCTGTTTTTCAGGTGGCGCTCTCACAAAGTCTTCAGGAAGTGAATCTGCCT	166287
QY 205	IleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArgProIleSerGly 224	
Db 166286	ATTTGACCTGGAGAGTGTGGCGAGCTCTGTTAACTAAGAGGCCATCAGTGGG 166227	
QY 225	LysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys-----	240
Db 166226	AAGACCTTTCTTGACACAGTTTCTCTGATGGGGGAGACGCATGTCAGGTGGGTGGG	166167
QY 240	-----	240
Db 166166	AGCTGGGGGCTGTCAGGCCCCAGGCAAGGATTCCCTGTTTGTGTTTCTTCTAGTTTTT	166107
QY 240	-----	240
Db 166106	CTTAATTTTCAGTGAATTTAATTTAATTAGAAAGCATATTGGAGGGGTAGATAAAGAA	166047
QY 240	-----	240
Db 166046	TAGTCTGAAGAGCTATAAGGTCAAATAGTAGTTAATGTGTGAGAAAAATGTTAGA	165987
QY 240	-----	240
Db 165986	TCCAAACCACTTCCCTAAGAGGCTACCTGTATTACAGGTGAGGGGCTGCAGGAAGACTGA	165927
QY 240	-----	240

QY 240 ----- 240

Db 165926 TTCCAGGTAAGCAGGAATCTGGAAAAGAGGTTAGAGTACCAAGGAGGTTTACATCAA 165867

QY 240 ----- 240

Db 165866 AATAAAGGCAGCTGGAAAACAATCTCTTTTAGCAGGAAGTGTGATTTAAAGAGGCTCTT 165807

QY 240 ----- 240

Db 165806 CTGCTTTTCCAAAGACCACAAATAACAGAGATAATCCCAAGATGAACTCGGGGTCTCC 165747

QY 240 -----,240

Db 165746 CATTCCTGAAGATGGCTGACTCTTTCCTATTTAGAAATTTTCCAGTGTGTGGATGTGGG 165687

QY 240 ----- 240

Db 165686 AACTTTTACACTTTTCCCTCCATTCTTCTACTACCAAGGCCCTGAAGTCAGTCTCATTTGCT 165627

QY 240 ----- 240

Db 165626 GGCCAAGACTTCTCTGAGTCTGCAGCATGAGAGGAGAGAAAGAAAGTCTGCTGCTTT 165567

QY 240 ----- 240

Db 165566 GATGAAAGCCATGGATGAGGAAAAGGAGGTGGGATTTTGGGTTCCAGTGATAGTTGAG 165507

QY 240 ----- 240

Db 165506 GATCATATGGGAATTGTGCTGGGATCTGTATACACAGGACCCCTGAGTCTAAATTTATAGA 165447

QY 240 ----- 240

Db 165446 AGTGGAGATGTTCATTGATTCCTCCTCCAGCATCTTGTGAGGAAACAGTGAGCATTTTG 165387

QY 241 -----GlnGlyAspSerG1 245

Db 165386 GGGCATAGACACTGACAGTGTAGTCTAATAGAGGCTTTTCCCTCTCGAGGGAGATTCAGG 165327

QY 245 yGlySerLeuMetCysArgAsnLysGlyAlaIleTrpThrLeuAlaGlyValThrSerTr 265

Db 165326 AGGTTTCACTCATGTGCGGAATAAGAAAGGGCCCTGACTCTGGCTGGTGTGACTTCCCTG 165267

QY 265 pGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSerAspGlnGlySerPr 285

Db 165266 GGGTTTGGGCTGTGGTGGAGGCTGGAGAAACAATGTGAGGAAAAGTCAAGGATCCCC 165207

QY 285 oGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrG1 305

Db 165206 TGGGATCTTTCACAGACATTTAGTAAAGTGTCTCCCTGGATCCACGAACACATCCAAACTGG 165147

QY 305 yAsn 306

Db 165146 TAAC 165143

RESULT 9

XL081291

LOCUS

DEFINITION

XP081291

ACCESSION

VERSION

U81291.1 GI:1754713

KEYWORDS

SOURCE

ORGANISM

Xenopus laevis.

Xenopus laevis

Xenopus laevis

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

Xenopodinae; Xenopus.

1 (bases 1 to 3689)

Lindsay, L.L., Wieduwilt, M.J., and Hedrick, J.L.

Oviductin, the Xenopus laevis oviductal protease that processes egg

envelope glycoprotein gp43, increases sperm binding to envelopes,

and is translated as part of an unusual mosaic protein composed of two protease and several CUB domains  
Biol. Reprod. 60 (4), 989-995 (1999)  
99184825  
PUBMED  
10084976  
2 (bases 1 to 3689)  
Yang, J.C., Lindsay, L.L. and Hedrick, J.L.  
Direct Submission  
Submitted (06-DEC-1996) Molecular & Cellular Biology, University of California, Davis, CA 95616, USA  
JOURNAL

## FEATURES

source

1..3689  
/organism="Xenopus laevis"  
/db\_xref="taxon:8355"  
7..3021  
/function="cleaves egg envelope glycoprotein gp43 to gp41  
rendering the egg fertilizable"  
/notes="oviduct serine protease"  
/codon\_start=1  
/product="oviductin"  
/protein\_id="AAB53972.1"  
/db\_xref="GI:1754714"  
/translation="MTRNLLLSILLSLAVKGDGPHRGARCGVPLGSATLNLVLS  
RIVGRESKQGPWTVSLRNKRGHFCGGTLVSHCHVLTAAHCLLDRNVKLYMRVYIG  
EYDQILKEETEQHVRVTEIFKPHNFNQSQPMNDYDVALLDGVSFTDENIQPACLFPN  
DDVFEPDCLVTILGWHLTENGILPVVQLQVLPIDLSCLHVMGSAKLTGVVSSYIV  
CAGFPEGKDACOGDGGPLLCORRHGSVYHLGLTSMGCGRSKNNVFLPHNRKGS  
PGIETDIQKLLIAGVNSQNTAVPNKNOESCQMODGVLSCKSGELIFLKNPMSVTRTMS  
GAGFSLSLCTCTSCNLFTHLDIIESFACNDLXLAITYDSHRLIGRFGCDIPRSLLI  
SFSSILNFFDHEHNTGFVLYISAVENPTYPSDSCGSPFVLFEEGEGISQPMYPNY  
LNSRCHWIIHSPSGSYIKIQDFEALFEDDDCSYLAIVYDAAEDKIEETFCGFSL  
PAPVSYTAVMHKIFSTDERDNDKGFRAFTFTFVPSNLVSDSRQGNMPTNKKETTAQ  
DSCIGVSQVPIPIFYNSIAKVEEAPHSWPHWTSLQYAGBHVCDGAILAPNWLITTA  
CVLNRRNDVWLDGILHDLRPHNKGKLVKQIIPHSFSDNDFDIALVELDSL  
QFNSDIPICLPKTBELAPASLCVYSGWSLRSKEAEKSTKLQOREVPIITLDACSAH  
YIQNGGITDMLCAGIGTQDNDSCSEQSGPLVLEKKGIYTFIGIASWGNCKE  
NSKPGIYKVSPIQWIRQIMSDTQISHLNGLPKPHPMGNIETETAGRDITQGGFP  
TNDASNNQNIYIASSCEDVVLQSGEIKMETKQMPNGFSCQWRIIAPKFQIILV  
MKQVHMSAENGKCCNSLIYIEGISKNTKLVRPTDMVPTWSESGSSVITESPFP  
VDFEFGFLVYSFHSRTQSDHVVPSDSSEP"

misc\_feature  
142..897  
/note="encodes protease domain"  
misc\_feature  
922..1263  
/note="encodes CUB domain"  
misc\_feature  
1291..1629  
/note="encodes CUB domain"  
misc\_feature  
1756..2472  
/note="encodes protease domain"  
misc\_feature  
2620..2964  
/note="encodes CUB domain"

BASE COUNT 1074 a 812 c 844 g 959 t  
ORIGIN

Alignment Scores:  
Pred. No.: 4.46e-55 Length: 3689  
Score: 688.50 Matches: 132  
Percent Similarity: 60.33% Conservative: 49  
Best Local Similarity: 44.00% Mismatches: 110  
Query Match: 42.08% Indels: 9  
DB: 5 Gaps: 4

US-09-735-713A-2 (1-306) x XL081291 (1-3689)

Qy 8 SerArgAsnLysLeuIleLeuLeuGlyTleValPhePheGluArgGlyLysSerAla 27  
:::||||| ||||||||| |||||  
Db 13 ACACGGAC-----TTACTTCTAGGATCTATTCTTAAGCCCTGGCTGAAGGG 63  
Qy 28 AlaLeuSerLeuProLysAlaProSerCysGlyGlnSer----LeuValLysValGlnPro 46  
::: ||||| |||||  
Db 64 GATCCTGGACCCCATAGAGTGCACGGTGTGGCAGTAAAGCCCATCGGTTCTGCAACGAA 123  
Qy 47 TrpAsnTyrPheAsnIlePheSerArgIleLeuGlyGlySerGlnValGluLysGlySer 66  
||||| |||||||||:::||||| |||||

Db 124 TTGAACATAC-----CTTCTCGAATCTTTCGAGGAGAGAAATCAAGAAAGGTCAA 174  
Qy 67 TyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIleCysGlyGlySerIleVal 86  
:::||||| |||||||||::: ||||| |||||||:::|||||  
Db 175 CATCCATGGACAGTATCATTAACACGGAATGAAAGCATTTTTCGGGAACCCCTGTG 234  
Qy 87 SerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsnArgAsnIleValSerThr 106  
||| ||||| |||||||||::: |||||  
Db 235 AGCCATGTGTCATGACTACAGCTGCCACTGTTTACTCGACAGGAATGTGAATATAT 294  
Qy 107 LeuAsnValThrAlaGlyGlyTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeu 126  
::: ||| |||||||||::: |||||  
Db 295 ATGAGAGTTTACATTCGAGAAATATGCAAAATCTAAAGAAAGAACTGAACAGATGTC 354  
Qy 127 ThrIleGluThrValIleIleHisProHisPheSerThrLysLysPrometAspTyrAsp 146  
::: |||||:::|||||::: ||||| |||||||:::|||||  
Db 355 AGGTTTATCGAATCTTCAACACACCCCACTTAACACAGATCAGCCAACTAATGAT 414  
Qy 147 IleAlaLeuLysMetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCys 166  
:::|||||::: ||||| |||||  
Db 415 GTTGCTGTACTGTGCTGGATGTTCCGTAACCTTTGATGAGAATATTTCAACACGATGC 474  
Qy 167 LeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGly 186  
|||||::: ||||| ||||| |||||  
Db 475 CTGCCAACCCAGATGATGTTTGAACCCGGGTGACCTCTGTACTCTTGGGTGGGT 534  
Qy 187 ArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeu 206  
||||||| |||||::: ||||| ||||||||| |||||  
Db 535 CATTACAGAGAAATGGAATCTCCCTGTTGTTTACAGAGGTCTACTTACCATTGTG 594  
Qy 207 ThrTrpGluGluCysValAlaAlaLeuLeuThrLysLysArgProIleSerGlyLysThr 226  
|||||::: |||||  
Db 595 GACCTCTCATCGTCCTGCTGATGTCAGCTCTGAAAGGACCGTGTCTCTCGTAT 654  
Qy 227 PheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGly 246  
:::|||||:::|||||:::|||||:::|||||  
Db 655 ATAGTGTGTCAGGGTTTCCAGAGGAGGAGGACGCTGCCAGGGGACTCAGGAGGC 714  
Qy 247 SerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGly 266  
|||||::: ||||| ||||| |||||  
Db 715 CCGTTGCTGCCAGCGCGTCATGGAAGTGGGTCTGTCATGCTCACCTCTCTGGGG 774  
Qy 267 LeuGlyCysGlyArgGlyTrpArgAsnVal-----ArgLysSerAspGlnGlySer 284  
:::||||||| |||||:::|||||  
Db 775 ATGGGTGTGGACGCTTCATGGAATAAATGTGTTTTCCTCATATAAGGAAAGGATCT 834  
Qy 285 ProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThr 304  
||||| ||||||||| |||||:::||||| |||||  
Db 835 CCAGGCATTTTCACAGACATCCAGAAATGCTTGGCTGGGTGAGCTCCGACGTTGAATACA 894

## RESULT 10

AB070367

LOCUS

AB070367

DEFINITION

AB070367

ACCESSION

AB070367.1

VERSION

AB070367.1

KEYWORDS

GI:15277253

SOURCE

Bufo japonicus female oviductal pars recta cDNA to mRNA.

ORGANISM

Bufo japonicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Neobatrachia; Bufonoides; Bufonidae;

Bufo.

REFERENCE

1

AUTHORS

Hyoshi, M., Takamune, K., Mita, K., Kubo, H., Sugimoto, Y. and

Katairi, C.

TITLE

Oviductin, the oviductal protease that mediates gamete interaction

by affecting the vitelline coat in Bufo japonicus: its molecular

cloning and analyses of expression and posttranslational activation

Dev. Biol. 243 (1), 176-184 (2002)

JOURNAL

MEDLINE

21835675

REFERENCE

2

(bases 1 to 3028)

AUTHORS

Takamune, K. and Hyoshi, M.

TITLE

Direct Submission

**JOURNAL**

Qy	132	Ile	le	His	Pro	His	Phe	Ser	Thr	Lys	Lys	Pro	Met	Asp	Tyr	Asp	Ile	Ala	Leu	Leu	Leu	151
Db	460	TT	CA	AG	CAC	CCCA	CT	TC	AA	CC	AG	CAG	CAG	CC	CTT	CA	AT	CA	TG	AC	CT	519

QY 152 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171  
::: ::: ||| :::: ||| ||||| |||  
Db 520 CTGGTGGAAAGCATCATTTGTATAAGATATCCAACCGGCATGCGTACCGAGCCCGAT 579

**Qy** 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly 191  
::: ||| ||| ::||| ||||||||| |||  
**Db** 580 GATGTATTCCACAGGAACCTCTGTATGGCACTTGGTTGGGCGCTTCGAGGAAAT 639

Qy 192 GlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTrpGluGluCys 211  
||| ||| |||||:|||| |||||:||||: ||| |||  
Db 640 GGAAGGTTACTAGACGCTGC AAAAGTTGTTTACCTCTATTGAGTACAGAAAGGTGC 699

Qy	212	ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly	231
		::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
		::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::	
Db	700	TTGAGCATCATGGAGACTGTGGCGTTTGGCACTTTGAGACTGTGGTTTGTGCAGGA	759

**Qy**      232    PheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg 251  
             |||||:::||:::||:::||:::||:::||:::||:::||:::||:::||:::||:::||:::||  
**Db**      760    TTTCAGAGAGGGGTAAAGACCGCTGTACGGTGACTCTCGAGGTCCTTTCTCTGCCAG 819

**Qy**    252 AsnLysLysGlyLaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGI.yArq 271  
               ::::| | | | | | | | | | | | | | | |  
**Db**    820 AGGAGCCACAGAGAGGTGGTGGTTGGTGCTAACTCCTGGGGATTAGCATCGCACGA 879

Qy 272 GlyTrpArgAsnValArgLys-----SerAspGlnClySerProGlyIlePheThr 289

Db 880 AATGGGCAGACAAATATCTTGGACCAGTGGAAGAAGCAAGGTTCACCGAGGTGTTTTACG 939

Qy 290 AspIleSerIysValIeuSerTrpIleHisGluHisIle 302  
 ||||| :|||:||||:||||: |||:||||:  
 Db 940 GATATTGAGGGCTTCTCAACTGGCTGAGTGAACCTC 978

RESULT 11
XLU81290
LOCUS
XLU81290
4628 bp
mRNA
linear
VPT 11-OCT
Xenopus laevis polyprotein, serine proteases and oyochninase
DEFINITION

regions, mRNA, complete cds.  
 U81290 U4951  
 U81290.1 GI:2981640  
 U81290.1 GI:2981640

SOURCE	ORGANISM
Xenopus laevis.	Xenopus laevis
Xenopus laevis	Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;	Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;

Xenopodinae; *Xenopus*.  
1 (bases 1 to 4628)

REFERENCE  
Lindsay, L.L., Yang, J.C. and Hedrick, J.L.  
Ovocytinase, a *Xenopus laevis* egg extracellular protease, is

JOURNAL  
MEDLINE  
PURWED 10500163  
translated as part of an unusual polyprotease  
Proc. Natl. Acad. Sci. U.S.A. 96 (20), 11253-11258 (1999)

2 (bases 1 to 4628)  
Yang, J.C., Lindsay, L.L. and Hedrick, J.L.  
cDNA Cloning of Ovocytochrome c, a Chymotrypsin-like Protease Released From Xenopus laevis Eggs at Fertilization

**JOURNAL**  
**REFERENCE**  
**AUTHORS**  
3 (bases 3737 to 4608)  
Yang, J.C. and Hedrick, J.L.  
Direct Submission

DIRECT SUBMISSION  
 JOURNAL  
 Submitted (29-DEC-1995) Molecular and Cellular Biology, Univer  
 of California at Davis, Davis, CA 95616, USA  
 REFERENCE  
 4 (bases 3736 to 4628)

AUTHORS Yang,J.C., Lindsay,L.L. and Hedrick,J.L.  
TITLE Direct Submission  
JOURNAL Submitted (06-DEC-1996) Molecular & Cellular Biology, University of California, Davis, CA 95616, USA

REFERENCE 5 (bases 1 to 4628)

AUTHORS Yang,J.C., Lindsay,L.L. and Hedrick,J.L.

JOURNAL Direct Submission

REMARK Nucleotide and amino acid sequence updated by submitter

COMMENT On or before Oct 11, 2000 this sequence version replaced

91:1216404, gi:1754711.

FEATURES

source

1..4628

/organism="Xenopus laevis"

/db\_xref="taxon:8355"

11..4585

/codon\_start=1

/product="polyprotein"

/protein\_id="AAC24717.1"

/db\_xref="GI:2981641"

/translation="MWYSITLLWVAACISAKISNLPPOESPPIGETAELKCGTRPE  
IGDEPDLEFTRIVGGDAAGVGPWVSLNERHCIGGSIVRKDWVTAARCVIPV  
TEIKVSHMTVIVSGYDQDMDSQESIPVSHIEPHNPYRGDNGMYDIALVFLSKPII  
FGSQVPTCLPQVGEKIEAGTLCVSSGMRLSENGDLSVQLQEVKLPVVDNMGTHAVL  
EPIGHVLDPTMLCAGFPFGMGADCGDSGGPFVCRRRSGVWFLAGCVSMGLGCRSW  
CAKQIIRSQSGPAIFSRVSVLDFLRPKLKTGGCSKGRITCKNGTVRYPLSGNYS  
INSYKRWMLAVOKAKTIEIRFLQIDIEDHATCTFDYLSFTVNEKMKIRKCGSTIPSL  
IVNSKRVTFPSDGTGTRGFEIQFLAIPKAAACSGSAKILKKMKIYSPNPDPY  
PLRTCSIIICAPENHFKAFEDFNEIGHGCIYDAVEYDGAEEKOLYIARLCGYTL  
PLPISSEPTMLIRFKTDMENSIFGKVFQSFVPEKQFSLPDDPTTISMLHPRAIA  
LDVCGMAPMTKMLPRIVGEEASPNSPMQVQIFLRTFHCAGIISQWILTAH  
CIRAAEPSTVYIAGDHNRLNTEQIRNIKTIRIHDYNSYTDNDIALLYLEEPL  
DLNDFVRVCLPEEVEVLTASVCTVWTGNGTADGQALGLOQLPILDSIICNTS  
YYSGLTDHMLCAGFPSSKEDACQDSGGLVCQNEKEQFSIYGLVSMGCGCRVSK  
PGVTVKELFTFQNTQDLOQENALLNSKVSQREGKAAGSGSSDVLVIELVGLIA  
TRGSLKSNKVYISQAQGVGHFGSGDGSNRNQGFFLQYIYVQVPDLAEFTGLSQ  
KPGYGSSECDVYISSEGVIESPNYLGNYPDLCWRITAPGVKTVIRSDVLELTKTEK  
DVSQCDPLLNGYIGESKDLGVCGEIRISLSKSGSEITLFTSVSEVSGKFSL  
IYSYKQKGLQEVLYLNLEASPNCSDYLNIDYGNPNSHLLASLCGHQODLVLSN  
GTFTVLHPHSDKSVVDWGFKIQYGNVPEPSTLASLDVPESEDDSEDESETKQKD  
CCGNTVDDPMLYMARSGKIRNLNKGKESGRVVGQQAAPRSPWLVSIOINSKRHY  
CGGIIITNKTLITAAHCEVKNLHRVVVGHDTLVEQNEHAKVYNHSHVHELYMPSGP  
PRNDLLLELPTLLNNSAVICLPDDVTDWTQAECLVAGWGVTDVGGMSLPTKLO  
QAKVPIVSTKCKDYWSDVTDNNICAGKAGSCMGDSGGGLICKREDRYLYLVGVVS  
WGSCKQDVKAPSVYLTLSAFMDWISQHMDT"

mat\_peptide

/product="serine protease"

mat\_peptide

/product="serine protease"

mat\_peptide

/product="ovochoymase"

/function="role in the blocking of polyspermy"  
/note="egg chymotrypsin-like protease; located in the  
perivitelline space of the unactivated egg and is released  
at fertilization"

misc\_feature

/note="encodes a CUB domain"

misc\_feature

/note="encodes a CUB domain"

misc\_feature

/note="encodes a CUB domain"

misc\_feature

/note="encodes a CUB domain"

misc\_feature

/note="encodes a CUB domain"

BASE COUNT

1257 a 1025 c 1166 g 1180 t

ORIGIN

Alignment Scores:

Pred. No.: 4.62e-49 Length: 4628

Score: 626.00 Matches: 128  
Percent Similarity: 57.89% Conservative: 37  
Best Local Similarity: 44.91% Mismatches: 110  
Query Match: 38.26% Indels: 10  
Gaps: 4

US-09-735-713A-2 (1-306) x XLU81290 (1-4628)

Qy 24 GlySerSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyClnSerLeuValLys 43

Db 104 GTGAAATGCGCAGAGCTCAAGTGTGGACACGCTCGTAAATTTGGTAT-----151

Qy 44 ValGlnProTrpAsnTrpPheAsnIlePheSerArgIleLeuGlyGlySerGlnValGlu 63

Db 152 ---GAACCAGACCTGGAATTCAC-----TCCCGCATTTGGGAGGAGAGATCTGTCT 202

Qy 64 LysGlySerTrpProTrpGlnValSerLeuLysGlnArgGlnLysHisIleCysGlyGly 83

Db 203 GTTGGAGGACCAACCATGGACGCTCGCTAAAGCTGAATGAGAGGCATATCTGTGGGGC 262

Qy 84 SerIleValSerProGlnTrpValIleThrAlaAlaHisCysIle-----AlaAsnArg 101

Db 263 AGCATTTGTCGCGAAGACATGTTGTAATCGCGGCCACTGTGTCTATCTCTGTGACTGAG 322

Qy 102 AsnIleValSerThrLeuAsnValThrAlaGlyGluTytrAspLeuSerGlnThrAspPro 121

Db 323 ATAAAGTGAGCCACATGATCTGATTGTTGGGAATATGACCAGCAAGTTATGGACACT 382

Qy 122 GlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLysLys 141

Db 383 CAAGACGAGAGATCCCGTATCCCATATAGAGCGCCACCAANTATACAGGGGATGGG 442

Qy 142 ProMetAspTytrAspIleAlaLeuLysMetAlaGlyAlaPheGlnPheGlyHisPhe 161

Db 443 ACATGGGCTATGACATGCTCTTCTCTTCTTAAGCCCAATCATCTTTGGCTCCAG 502

Qy 162 ValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThr 181

Db 503 GTGCACCCATTTGGCTTCCCGAGGTGGGAGAGAAGATTGAAGCCGGAACCTTGTGT 562

Qy 182 ThrAlaGlyTrpGlyArgLeuThrGluGlyValLeuSerGlnValLeuGlnVal 201

Db 563 TCCACGGCTGGGGAAGGCTTGAAGAAATGGGACTTGTCCCAAGTCTCATAGGAAGTA 622

Qy 202 AsnLeuProIleLeuThrTrpGluGluCysValAlaLeuLeuThrLysLysArgPro 221

Db 623 AAATGCCAGTCGTTGATAATGSCACCTGCCATGCTGTGTGGAGCCCATAGGCATCCC 682

Qy 222 IleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyArgAspAlaCysGln 241

Db 683 GTATTAGATGACACAATGTTATGTGCTGATTTCCAGAGAGGAGGATGATGCTGCCAG 742

Qy 242 GlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaThrPheLeuAlaGly 261

Db 743 GCGCATTCGAGAGACCTTTGTATGACAGCGCGCTCTGGGGTGTGGTCTTGGCTGGA 802

Qy 262 ValThrSerTrpGlyLeuGlyCysGlyArgGlyTrp---ArgAsnAsnValArgLysSer 280

Db 803 TGTGTATCTCTGGGCGCTGGCTGTGGGAGAAGCTGGGGTGGCCAAACAGATACACGACC 862

Qy 281 AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu 300

Db 863 CAGAGTGGATCCCTGCAATTTTCTCCAGGGTCTCTCAGTGTGGACTTCCTGAGACCC 922

Qy 301 HisIleGlnThrGly 305

Db 923 CCTAAACTAACTGGA 937

RESULT 12

AX480940

LOCUS AX480940 2627 bp DNA linear PAT 12-AUG-2002

DEFINITION Sequence 36 from Patent WO0246383.

ACCESSION AX480940



Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.	26458	28550:	contig of 2093 bp in length
Unpublished	28551	28650:	gap of unknown length
2 (bases 1 to 169388)	28651	29929:	contig of 1279 bp in length
Direct Submission	29930	30029:	gap of unknown length
Worley,K.C.	30030	31427:	contig of 1398 bp in length
Submitted (31-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	31428	31527:	gap of unknown length
----- Genome Center	31528	33329:	contig of 1802 bp in length
Center: Baylor College of Medicine	33330	33429:	gap of unknown length
Center code: BCM	33430	34854:	contig of 1425 bp in length
Web site: <a href="http://www.hgsc.bcm.tmc.edu/">http://www.hgsc.bcm.tmc.edu/</a>	34855	34954:	gap of unknown length
Contact: hgsc-help@bcm.tmc.edu	34955	36498:	contig of 1544 bp in length
----- Project Information	36499	36598:	gap of unknown length
Center project name: KAHL	36599	39161:	contig of 2563 bp in length
Center clone name: CH230-388116	39162	39261:	gap of unknown length
----- Summary Statistics	39262	40905:	contig of 1644 bp in length
Sequencing vector: Plasmid	40906	41005:	gap of unknown length
Chemistry: Dye-terminator Big Dye: 100% of reads	41006	44218:	contig of 3213 bp in length
Assembly program: Phrap; version 0.990329	44219	44318:	gap of unknown length
Consensus quality: 136807 bases at least Q40	44319	46887:	contig of 2569 bp in length
Consensus quality: 141381 bases at least Q30	46888	46987:	gap of unknown length
Consensus quality: 145526 bases at least Q20	46988	49280:	contig of 2293 bp in length
-----	49281	49380:	gap of unknown length
* NOTE: Estimated insert size may differ from sequence length	49381	51242:	contig of 1862 bp in length
*(see <a href="http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html">http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html</a> ).	51243	51342:	gap of unknown length
* NOTE: This is a 'working draft' sequence. It currently	51343	53327:	contig of 1985 bp in length
* consists of 56 contigs. The true order of the pieces	53328	53427:	gap of unknown length
* is not known and their order in this sequence record is	53428	55718:	contig of 2291 bp in length
* arbitrary. Gaps between the contigs are represented as	55719	55818:	gap of unknown length
* runs of N, but the exact sizes of the gaps are unknown.	55819	59313:	contig of 3495 bp in length
* This record will be updated with the finished sequence	59314	59413:	gap of unknown length
* as soon as it is available and the accession number will	59414	61947:	contig of 2534 bp in length
* be preserved.	61948	62047:	gap of unknown length
1	62048	64296:	contig of 2249 bp in length
1003	64297	64396:	gap of unknown length
1103	64397	66012:	contig of 1616 bp in length
2133:	66013	66112:	gap of unknown length
2233:	66113	68206:	contig of 2094 bp in length
3358:	68207	68306:	gap of unknown length
3458:	68307	71222:	contig of 2916 bp in length
5023:	71223	71322:	gap of unknown length
5123:	71323	74509:	contig of 3187 bp in length
5124	74510	74609:	gap of unknown length
5124	74610	77284:	contig of 2675 bp in length
6684	77285	77384:	gap of unknown length
6784	77385	80659:	contig of 3275 bp in length
8066	80660	80759:	gap of unknown length
8166	80760	84060:	contig of 3301 bp in length
9312:	84061	84160:	gap of unknown length
9413	84161	86617:	contig of 2457 bp in length
11980:	86618	86717:	gap of unknown length
12080:	86718	89029:	contig of 2312 bp in length
13407:	89030	89129:	gap of unknown length
13507:	89130	92826:	contig of 3697 bp in length
13508	92827	92926:	gap of unknown length
13508	92927	97425:	contig of 4499 bp in length
13508	97426	97525:	gap of unknown length
15457:	97526	100539:	contig of 3014 bp in length
15557:	100540	100639:	gap of unknown length
15558	100640	106363:	contig of 5724 bp in length
16567:	106364	106463:	gap of unknown length
16667:	106464	111809:	contig of 5346 bp in length
16668	111810	111909:	gap of unknown length
17971:	111910	115701:	contig of 3792 bp in length
18071:	115702	115801:	gap of unknown length
19740:	115802	121844:	contig of 6043 bp in length
19740:	121845	121944:	gap of unknown length
19841	121945	130099:	contig of 8155 bp in length
21943:	130100	130199:	gap of unknown length
22043	130200	137395:	contig of 7196 bp in length
23410:	137396	137495:	gap of unknown length
23510:	137496	143948:	contig of 6453 bp in length
25146:	143949	144048:	gap of unknown length
25245:	144049	151651:	contig of 7603 bp in length
25147			
25247			
26357:			
26457:			
26358			

Db 98598 TGATGAGCCTAAAGAAATGGGAAAAAGACGACGCAATTGCACAGATTCTGATGTTGGGGATC 98657

DB:	3	Gaps:	12
US-09-735-713A-2 (1-306) x SSU79521 (1-942)			
QY	10	AsnLysLeuIleLeuLeuGlyIleValPheGluArgGlyLysSerAlaAlaLeu	29
DB	24	ATACGCTTCACATCTTATC---GTGACATATATTT	56
QY	30	SerLeuProLysAlaProSerCysGlyGlnSerLeuValLysValGlnProTrpAsnTyr	49
DB	57	TCACATTGATTTGGTCCAGATGTGA-----ATMAAAACGGACCT-----ATG	101
QY	50	PheAsnIlePheSerArgIleLeuGlyCysSerGlnValGluLysGlySerTyrProTrp	69
DB	102	TTAGTAGAGTTCAATCGAATCGTGGTGGAGAACCTGCCAACCAGGAGATTTCCCTGG	161
QY	70	GlnValSerLeuLys-----GlnArgGlnLysHisIleCysGlySerIle	85
DB	162	CAGATATCATTTACAAAGTCGTCTGTGTATGGAAGTTACCATATTTCGGTGGATCATTT	221
QY	86	ValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsnArgAsnIleValSer	105
DB	222	CTCGATGAAGATTGGTGTCTGCTGCTCATTTGTGAGAGGGATGAAT---CCTAGT	278
QY	106	ThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThr	125
DB	279	GATCTTAGGATTTCTGAGGAGACATACTTCAAAAAGAGCGCATGACAAATGG	338
QY	126	LeuThrIleGluThrValIleIleHisProHisPheSerThrLysLysProMetAspTyr	145
DB	339	CAAGATGTAATATGATCATCATCATCAAGAAGATTAC---GTTTATTCTACTCTTGAAC	395
QY	146	AspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPhe---GlyHisPheValGlyPro	164
DB	396	GATATAGCCCTTTTGAATATGGCCGAACCTTTGGATTGTACTCCGACTGCGATTGGATCC	455
QY	165	IleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThrAlaGly	184
DB	456	ATTTGTTTGCCTAGTCAAAATAATCAGAATTTTCTGGCCAT---TGTATAGTAACCGGA	512
QY	185	TrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnValAlaAsnLeuPro	204
DB	513	TGGGATCTGTGCAGAGAAGTGTGAATTCCTCCCAACATATTACAAAAGTTTCTGTCCA	572
QY	205	IleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArgProIleSerGly	224
DB	573	TTAATGACCGATGAAGATGCGAGTAATCTACAATATAGTG-----	614
QY	225	LysThrPheLeuCysThrGlyPheProAspGlyArgAspAlaCysGlnGlyAspSer	244
DB	615	GATACCATGTTATCTGCAGATATGCAGAGAGGGGCAAAAGTCTGTCTCAGGTGATCA	674
QY	245	GlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSer	264
DB	675	GGTGGTCCCTTGGTCTGCCAACGGAGACGAACCTATTCTTGGCCGGGATGTATCC	734
QY	265	TrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSerAspGlnGlySer	284
DB	735	TGGGGCATCGGATGTCTCAACCT---CGTAAT-----	764
QY	285	ProGlyIlePheThrAspIleSerLysValLeuSerTrpIle	298
DB	765	CCTGGATTTACACACAAGTATCAAAATTTCTTGATTGGATA	806
RESULT 15			
AX360098			
LOCUS			
Sequence 54 from Patent WO0200860.			
AX360098			
ACCESSION			
VERSION			
KEYWORDS			
human.			
ORGANISM			
Homo sapiens			

Qy 289 ThrAspIleSerLysValLeuSerTrpIleHisGlu 300  
||| |||::: |||:::||||||| :::  
Db 2362 ACCCGCATCACAGGTGTGATCAGCTGGATCCAGCAA 2397

Search completed: March 25, 2003, 03:02:41  
Job time : 3507 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 02:04:19 ; Search time 50 Seconds  
(without alignments)  
815.494 Million cell updates/sec

Title: US-09-735-713A-2  
Perfect score: 1636  
Sequence: 1 MSLKMLISRNKLLILLGIVF.....IPTDISKVLSMIHEIQTGN 306

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002:\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*

2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*

3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*

4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*

5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*

6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*

7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*

8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*

9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*

10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*

11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*

12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*

13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*

14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*

15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*

16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*

17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*

18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*

19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*

20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*

21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*

22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*

23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1636	100.0	306	AAE04733	Human protease hom
2	1632	99.8	306	AAU074760	Human protease PRT
3	1618	98.9	302	AAE04734	Human protease hom
4	1362	83.3	556	AAU82748	Amino acid sequenc
5	950	58.1	182	AAU24275	Toad EST encoded p
6	898	54.9	164	AAE04735	Human protease hom
7	567.5	34.7	913	ABG24246	Novel human diagno
8	567.5	34.7	1576	ABG07870	Novel human diagno
9	567.5	34.7	1576	ABG10218	Novel human diagno
10	567.5	34.7	1576	ABG14588	Novel human diagno

11	567.5	34.7	1576	22	ABG19887	Novel human diagno
12	561.5	34.3	1031	23	ABP60993	Novel human protei
13	489.5	29.9	802	20	AAU41710	Human PRO618 prote
14	489.5	29.9	802	21	ABAB4266	Human PRO618 (UNQ3
15	489.5	29.9	802	21	AAU24052	Human PRO618 prote
16	489.5	29.9	802	23	AAU82755	Amino acid sequenc
17	486	29.7	787	22	ABU71302	Drosophila melanog
18	485.5	29.7	658	22	AAE06934	Human membrane-ty
19	485.5	29.7	802	22	AAE06933	Human membrane-ty
20	476.5	29.1	235	22	AAE06932	Human membrane-ty
21	467	28.5	452	20	AAU41694	Human PRO382 prote
22	467	28.5	453	22	AAU29055	Human PRO polypept
23	467	28.5	453	22	AAE06935	Human membrane-ty
24	467	28.5	453	23	AAE23020	Human trypsin fami
25	467	28.5	454	21	ABAB2246	Tumour associated
26	467	28.5	454	22	AAU68911	Human TAGB-12, pro
27	466	28.5	453	21	ABAB4250	Human PRO382 (UNQ3
28	456.5	27.9	327	21	AAU72093	Human serine prote
29	456.5	27.9	327	23	ABAB17921	Human gene 3 encod
30	456.5	27.9	394	23	ABP41994	Human ovarian anti
31	456.5	27.9	454	23	AAE23024	Human trypsin fami
32	456.5	27.9	454	23	AAU82745	Amino acid sequenc
33	454.5	27.8	248	21	ABAB3572	Human cancer assoc
34	454.5	27.8	1128	23	AAU98890	Human protease PRT
35	454.5	27.8	1128	23	AAU82739	Amino acid sequenc
36	446	27.3	243	23	AAU80516	Epithelin-like ser
37	446	27.3	309	23	AAU80531	Epithelin-like ser
38	445	27.2	818	23	AAU82753	Amino acid sequenc
39	444.5	27.2	328	21	ABAB36480	Fusion gene with h
40	444.5	27.2	328	22	ABAB67540	Amino acid sequenc
41	444	27.1	343	23	AAU78547	Human prostaticin pr
42	444	27.1	343	23	ABU07285	Amino acid sequenc
43	444	27.1	414	21	ABAB08912	Human secreted pro
44	444	27.1	480	21	ABAB08950	Human secreted pro
45	442.5	27.0	417	22	AAE06942	Human hepsin prote

ALIGNMENTS

RESULT 1

AAE04733

ID AAE04733 standard; Protein; 306 AA.

XX

AC AAE04733;

DT 10-SEP-2001 (first entry)

XX

Human protease homologue #1.

XX

Human; protease homologue; novel human protein; NHP; therapy; pharmacogenomic application; physiological disorder.

XX

Homo sapiens.

XX

Key Location/Qualifiers

FT Misc-difference 23 /note= "Encoded by CRA"

FT Misc-difference 28 /note= "Encoded by RCT"

XX

WO200146407-A1.

XX

28-JUN-2001.

XX

12-DEC-2000; 2000WO-US33738.

XX

23-DEC-1999; 99US-0171566.

XX

(LEXI-) LEXICON GENETICS INC.

XX

Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;

DR WPI; 2001-408641743.  
DR N-PSDB; AAD09328.  
XX  
PT Polynucleotide encoding novel human protease homologs, useful for  
PT identifying agonist, antagonist or modifiers or for producing  
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic  
PT applications -  
XX  
PS Claim 2; Page 28-29; 32pp; English.  
XX  
CC The present sequence is novel human protein (NHP).  
CC known as human protease homologue. NHP shares structural similarity  
CC with animal proteases, particularly trypsin-like protease such  
CC as oviductin, plasminogen activator and human plasma kallikrein  
CC precursor. NHP and its cDNA are useful in therapeutic, diagnostic and  
CC pharmacogenomic applications. NHP sequences are useful for identifying  
CC agonists, antagonists and modulators and also for producing antibodies  
CC useful in diagnosis, drug screening, clinical trial monitoring and in  
CC treatment of physiological disorders.  
XX  
SQ Sequence 306 AA;  
Query Match 100.0%; Score 1636; DB 22; Length 306;  
Best Local Similarity 100.0%; Pred. No. 2.3e-130;  
Matches 306; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MSLKMLISRNKLLILLGIVFFERKSAALSLPKAPSCQSLVKVQPMWNYNIFSRILGGS 60  
Db 1 MSLKMLISRNKLLILLGIVFFERKSAALSLPKAPSCQSLVKVQPMWNYNIFSRILGGS 60  
Qy 61 QVEKGSYPWQVSLKQKHICGGSTVSPQWVITAAHCIAANRNIVSTLNTVAGEYDLSQTD 120  
Db 61 QVEKGSYPWQVSLKQKHICGGSTVSPQWVITAAHCIAANRNIVSTLNTVAGEYDLSQTD 120  
Qy 121 PGEQTLTIETVLIHPHFSTKKPMDDYDIALKMGAFQFGHFVGPICLPRLREQFEAGFIC 180  
Db 121 PGEQTLTIETVLIHPHFSTKKPMDDYDIALKMGAFQFGHFVGPICLPRLREQFEAGFIC 180  
Qy 181 TTAGGRLTGEGVLSQVLOEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGRDAC 240  
Db 181 TTAGGRLTGEGVLSQVLOEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGRDAC 240  
Qy 241 QDGGSLMCRNKKGAWTLAGVTSWGLCGGRWRNNVRKSDQSGPIETDISKVLWIHE 300  
Db 241 QDGGSLMCRNKKGAWTLAGVTSWGLCGGRWRNNVRKSDQSGPIETDISKVLWIHE 300  
Qy 301 HIQTGN 306  
Db 301 HIQTGN 306  
RESULT 2  
AAU74760  
ID AAU74760 standard; Protein; 306 AA.  
XX  
AC AAU74760;  
XX  
DT 09-APR-2002 (first entry)  
XX  
DE Human protease PRTS-20 protein sequence.  
XX  
KW Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;  
KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;  
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;  
KW cell proliferative disorder; developmental disorder; epilepsy;  
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;  
KW reproductive disorder; endometriosis.  
XX  
OS Homo sapiens.  
XX  
PN W0200198468-A2.  
XX  
PD 27-DEC-2001.

XX 13-JUN-2001; 2001WO-US191178.  
XX  
XX 16-JUN-2000; 2000US-212336P.  
PR 22-JUN-2000; 2000US-213955P.  
PR 29-JUN-2000; 2000US-215396P.  
PR 07-JUL-2000; 2000US-216821P.  
PR 14-JUL-2000; 2000US-218946P.  
XX  
XX (INCY-) INCYTE GENOMICS INC.  
XX  
XX Yue H, Elliott VS, Gandhi AR, Lal P, Au-young J, Tribouley CM;  
PI Deleage AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;  
PI Walia NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;  
PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;  
PI Kallick DA;  
XX  
DR WPI; 2002-090437/12.  
DR N-PSDB; ABK12903.  
XX  
XX Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful  
PT in the diagnosis, treatment and prevention of gastrointestinal (e.g.  
PT gastritis), cardiovascular (e.g. atherosclerosis) and cell  
PT proliferative (e.g. cancer) disorders -  
XX  
PS Claim 1; Page 157-158; 177pp; English.  
XX  
XX The present invention relates to twenty one new human proteases,  
CC referred to as PRTS-1 to PRTS-21. The PRTS polynucleotides and  
CC polypeptides of the invention are useful in the diagnosis, treatment and  
CC prevention of gastrointestinal e.g. gastritis, esophageal carcinoma and  
CC Crohn's disease, cardiovascular e.g. atherosclerosis, hypertension and  
CC myocardial infarction, autoimmune/inflammatory e.g. acquired  
CC immunodeficiency syndrome (AIDS), allergies and osteoarthritis, cell  
CC proliferative e.g. cancer, developmental e.g. Duchenne and Becker  
CC muscular dystrophy, epithelial e.g. dermatitis, neurological e.g.  
CC epilepsy and Alzheimer's disease and reproductive e.g. infertility and  
CC endometriosis disorders. Numerous other examples of each disorder are  
CC given in the specification. The present protein sequence represents  
CC the human protease PRTS-20 protein of the invention.  
XX  
SQ Sequence 306 AA;  
Query Match 99.8%; Score 1632; DB 23; Length 306;  
Best Local Similarity 99.7%; Pred. No. 5.1e-130;  
Matches 305; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MSLKMLISRNKLLILLGIVFFERKSAALSLPKAPSCQSLVKVQPMWNYNIFSRILGGS 60  
Db 1 MSLKMLISRNKLLILLGIVFFERKSAALSLPKAPSCQSLVKVQPMWNYNIFSRILGGS 60  
Qy 61 QVEKGSYPWQVSLKQKHICGGSTVSPQWVITAAHCIAANRNIVSTLNTVAGEYDLSQTD 120  
Db 61 QVEKGSYPWQVSLKQKHICGGSTVSPQWVITAAHCIAANRNIVSTLNTVAGEYDLSQTD 120  
Qy 121 PGEQTLTIETVLIHPHFSTKKPMDDYDIALKMGAFQFGHFVGPICLPRLREQFEAGFIC 180  
Db 121 PGEQTLTIETVLIHPHFSTKKPMDDYDIALKMGAFQFGHFVGPICLPRLREQFEAGFIC 180  
Qy 181 TTAGGRLTGEGVLSQVLOEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGRDAC 240  
Db 181 TTAGGRLTGEGVLSQVLOEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGRDAC 240  
Qy 241 QDGGSLMCRNKKGAWTLAGVTSWGLCGGRWRNNVRKSDQSGPIETDISKVLWIHE 300  
Db 241 QDGGSLMCRNKKGAWTLAGVTSWGLCGGRWRNNVRKSDQSGPIETDISKVLWIHE 300  
Qy 301 HIQTGN 306  
Db 301 HIQTGN 306  
RESULT 3

AAE04734		
ID	AAE04734	standard; Protein; 302 AA.
XX	AAE04734;	
AC	AAE04734;	
XX		
DT	10-SEP-2001	(first entry)
XX		
DE	Human	protease homologue #2.
XX		
KW	Human;	protease homologue; novel human protein; NHP; therapy;
KW	pharmacogenomic	application; physiological disorder.
XX		
OS	Homo sapiens.	
XX		
XX	Key	Location/Qualifiers
FT	Misc-difference 19	/note= "Encoded by CRA"
FT	Misc-difference 24	/note= "Encoded by RCT"
FT		
XX	WO200146407-A1.	
PN		
XX	28-JUN-2001.	
PD		
XX		
XX	12-DEC-2000;	2000WO-US33738.
PF		
XX	23-DEC-1999;	99US-0171566.
PR		
XX	(LEXI-)	LEXICON GENETICS INC.
PA		
XX	Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;	
PI		
XX	WPI: 2001-408641/43.	
DR	N-PSDB; AAD09329.	
DR		
XX		
PT	Polynucleotide encoding novel human protease homologs, useful for	
PT	identifying agonist, antagonist or modifiers or for producing	
PT	antibodies useful in therapeutic, diagnostic and pharmacogenomic	
PT	applications -	
XX		
XX	Disclosure; Page 29-30; 32pp; English.	
PS		
XX	The present sequence is novel human protein (NHP),	
CC	known as human protease homologue. NHP shares structural similarity	
CC	with animal proteases, particularly trypsin-like protease such	
CC	as oviductin, plasminogen activator and human plasma kallikrein	
CC	precursor. NHP and its cDNA are useful in therapeutic, diagnostic and	
CC	pharmacogenomic applications. NHP sequences are useful for identifying	
CC	agonists, antagonists and modulators and also for producing antibodies	
CC	useful in diagnosis, drug screening, clinical trial monitoring and in	
CC	treatment of physiological disorders.	
XX		
SQ	Sequence. 302 AA;	
Query Match		98.9%; Score 1618; DB 22; Length 302;
Best Local Similarity		100.0%; Pred. No. 7.7e-129;
Matches 302; Conservative		0; Mismatches 0; Indels 0; Gaps 0;
QY	5	MLISRNKLIILLGIVFFERKSAALSIPKAPSCGQSLVKVQPNWYFNIFSRILGGSQVEK 64
Db	1	MLISRNKLIILLGIVFFERKSAALSIPKAPSCGQSLVKVQPNWYFNIFSRILGGSQVEK 60
QY	65	GSYPQWVSLKORQHICGGISVSPWVITAACHIANRNIVSTLNVTAGEYDLSOTDPGEQ 124
Db	61	GSYPQWVSLKORQHICGGISVSPWVITAACHIANRNIVSTLNVTAGEYDLSOTDPGEQ 120
QY	125	TLTETVTHPHFSTKKPMDYDIALKMGAFQFGHVGPICLPELREQFEAGFICTTAG 184
Db	121	TLTETVTHPHFSTKKPMDYDIALKMGAFQFGHVGPICLPELREQFEAGFICTTAG 180
QY	185	WGRITGGVLSQVLQEVNLPILTWEECVAALLTLKRPISGKTFICTGFPDGGRDACQGS 244
Db	181	WGRITGGVLSQVLQEVNLPILTWEECVAALLTLKRPISGKTFICTGFPDGGRDACQGS 240
Query Match		83.3%; Score 1362; DB 23; Length 556;

QY	245	GGSLMCRNKKGAWTLAGVTSWGLCCGRGWRNNVRKSDQSGPIFTDISKVLSWIHEHIQT 304
Db	241	GGSLMCRNKKGAWTLAGVTSWGLCCGRGWRNNVRKSDQSGPIFTDISKVLSWIHEHIQT 300
QY	305	GN 306
Db	301	GN 302
RESULT 4		
AAU82748		
ID	AAU82748	standard; Protein; 556 AA.
XX	AAU82748;	
DT	23-APR-2002	(first entry)
XX		
DE	Amino acid sequence of novel human protease #47.	
XX		
KW	Human; protease; cancer; immune-related disorder; cardiovascular disease;	
KW	neurological-associated disease; metabolic disorder; inflammatory disorder;	
KW	nervous system disorder; sexual dysfunction; pain; mood disorder;	
KW	hypertension; psychotic disorder; neurological disorder; dyskinesia;	
KW	viral infection; human immunodeficiency virus; HIV; non-viral infection;	
KW	ocular disease; cytostatic; enzyme.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200200860-A2.	
XX		
PD	03-JAN-2002.	
XX		
PF	26-JUN-2001; 2001WO-US20171.	
XX		
PR	26-JUN-2000; 2000US-214047P.	
XX		
PA	(SUGE-) SUGEN INC.	
XX		
PI	Plowman G, Whyte D, Sudarsanam S, Manning G, Caenepeel S;	
PI	Charyczak G;	
XX		
DR	WPI: 2002-139913/18.	
DR	N-PSDB; ABK31790.	
XX		
PT	Nucleic acids encoding novel human proteases, useful for useful for	
PT	treating diseases and disorders such as cancers, immune-related	
PT	diseases and disorders, cardiovascular disease (e.g. restenosis) and	
PT	inflammatory disorders -	
XX		
PS	Claim 6; Fig 2P; 313pp; English.	
XX		
CC	The present invention relates to the isolation of novel human	
CC	proteases, and the nucleic acids encoding them. The sequences of	
CC	the invention are useful for treating diseases and disorders such as	
CC	cancers (e.g. breast, colon, lung), immune-related diseases and disorders	
CC	(e.g. inflammatory diseases and asthma), cardiovascular diseases	
CC	(e.g. restenosis and coronary thrombosis), brain or neuronal-associated	
CC	diseases, metabolic disorders (e.g. diabetes, obesity), inflammatory	
CC	disorders (e.g. rheumatoid arthritis and psoriasis), central or	
CC	peripheral nervous system diseases, migraines, pain, sexual dysfunction,	
CC	mood disorders, attention disorders, cognition disorders, hypotension,	
CC	hypertension, psychotic disorders, neurological disorders	
CC	(e.g. Alzheimer's disease, Parkinson's disease) and dyskinesias.	
CC	The nucleic acids and polypeptides are also useful for treating viral	
CC	infections caused by human immunodeficiency virus (HIV), and non-viral	
CC	infections such as ocular disease (e.g. glaucoma) and macular	
CC	degeneration. AAU82702-AAU82760 represent the novel human proteases of	
CC	the invention.	
XX		
SQ	Sequence 556 AA;	

Best Local Similarity 85.7%; Pred. No. 6.7e-107;  
Matches 263; Conservative 8; Mismatches 28; Indels 8; Gaps 2;  
Qy 1 MSLKMLISRNKLIILLGIVFFERKSAALSLPKAPSCGSLVKVQPNWYFNIFSRILGGS 60  
Dy 1 MSLKMLISRNKLIILLGIVFFERKSAATLSLPKAPSCGSLVKVQPNWYFNIFSRILGGS 60  
Qy 61 QVEKGSYPQWVSLKQKHICGGISVSPQWVITAHCIAANRNIVSTLNVTAAGEYDLSQTD 120  
Dy 61 QVEKGSYPQWVSLKQKHICGGISVSPQWVITAHCIAANRNIVSTLNVTAAGEYDLSQTD 120  
Qy 121 PGEQTLTIETVIIHPHFSTKKPMYDIALLLKMGAFQGHFVGPICLPDLPELREQFEAGFIC 180  
Dy 121 PGEQTLTIETVIIHPHFSTKKPMYDIALLLKMGAFQGHFVGPICLPDLPELREQFEAGFIC 180  
Qy 181 TTAGHGRTEGVLSQLQEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGGDRAC 240  
Dy 181 TTAGHGRTEGVLSQLQEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGGDRAC 240  
Qy 241 QDGSGLMCRNKKGAWTLAGVTSMLGCGRWNRNVRKSDQSP-----GIFTDISK 293  
Dy 241 QDGSGLMCRNKKGAWD-SGWSIWEAQVGGSLSSRSRPSLGNKVRCLTNFNFFKLAG 299  
Qy 294 VLSWIHE 300  
Dy 300 CGTWCSE 306

RESULT 5  
AAM24275  
ID AAM24275 standard; Protein: 182 AA.  
XX AC AAM24275;  
XX DT 12-OCT-2001 (first entry)  
XX DE Toad EST encoded protein SEQ ID NO: 1800.  
XX KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
KW diagnostics; forensic test; gene mapping; genetic disorder;  
KW biodiversity; gene therapy; nutrition.  
XX OS Xenopus laevis.  
XX PN WO200154477-A2.  
XX PD 02-AUG-2001.  
XX PF 25-JAN-2001; 2001WO-US02687.  
XX PR 25-JAN-2000; 2000US-0491404.  
PR 17-JUL-2000; 2000US-0617746.  
PR 03-AUG-2000; 2000US-0631451.  
PR 15-SEP-2000; 2000US-0663870.  
XX (HYSE-) HYSEQ INC.  
XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
PI Cao Y, Drmanac RA, Zhang J, Werhman T;  
XX WPI; 2001-476164/51.  
DR N-PSDB; AAH98934.  
XX Isolated polypeptide for treatment of diseases, diagnostics, raising  
PT antibodies and research use -  
XX Claim 20; Page 1178-1179; 1275pp; English.  
XX The present invention provides the protein and coding sequences of novel  
CC proteins from a variety of organisms, including human, dog, cat, horse,  
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)

CC from the organism of interest. They can be used in diagnostics,  
CC forensics, gene mapping, identification of mutations, to assess  
CC biodiversity and for nutritional purposes. The present sequence is a  
CC protein of the invention.  
XX  
SQ Sequence 182 AA;  
Query Match 58.1%; Score 950; DB 22; Length 182;  
Best Local Similarity 99.5%; Pred. No. 1.3e-72;  
Matches 181; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MSLKMLISRNKLIILLGIVFFERKSAALSLPKAPSCGSLVKVQPNWYFNIFSRILGGS 60  
Dy 1 MSLKMLISRNKLIILLGIVFFERKSAATLSLPKAPSCGSLVKVQPNWYFNIFSRILGGS 60  
Qy 61 QVEKGSYPQWVSLKQKHICGGISVSPQWVITAHCIAANRNIVSTLNVTAAGEYDLSQTD 120  
Dy 61 QVEKGSYPQWVSLKQKHICGGISVSPQWVITAHCIAANRNIVSTLNVTAAGEYDLSQTD 120  
Qy 121 PGEQTLTIETVIIHPHFSTKKPMYDIALLLKMGAFQGHFVGPICLPDLPELREQFEAGFIC 180  
Dy 121 PGEQTLTIETVIIHPHFSTKKPMYDIALLLKMGAFQGHFVGPICLPDLPELREQFEAGFIC 180  
Qy 181 TT 182  
Dy 181 TT 182  
RESULT 6  
AAE04735  
ID AAE04735 standard; Protein: 164 AA.  
XX AC AAE04735;  
XX DT 10-SEP-2001 (first entry)  
XX DE Human protease, homologue #3.  
XX KW Human; protease homologue; novel human protein; NHP; therapy;  
KW pharmacogenomic application; physiological disorder.  
XX OS Homo sapiens.  
XX PN WO200146407-A1.  
XX PD 28-JUN-2001.  
XX PF 12-DEC-2000; 2000WO-US33738.  
XX PR 23-DEC-1999; 99US-0171566.  
XX (LEXI-) LEXICON GENETICS INC.  
XX Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;  
PI WPI; 2001-408641/43.  
DR N-PSDB; AAD09330.  
XX Polynucleotide encoding novel human protease homologs, useful for  
PT identifying agonist, antagonist or modifiers or for producing  
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic  
PT applications -  
XX Disclosure; Page 30-31; 32pp; English.  
XX The present sequence is novel human protein (NHP),  
CC known as human protease homologue. NHP shares structural similarity  
CC with animal proteases, particularly trypsin-like protease such  
CC as oviductin, plasminogen activator and human plasma kallikrein  
CC precursor. NHP and its cDNA are useful in therapeutic, diagnostic and  
CC pharmacogenomic applications. NHP sequences are useful for identifying  
CC agonists, antagonists and modulators and also for producing antibodies  
CC useful in diagnosis, drug screening, clinical trial monitoring and in



CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABC30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 1576 AA;

Query Match 34.7%; Score 567.5; DB 22; Length 1576;  
Best Local Similarity 41.2%; Pred. No. 3.3e-39;  
Matches 114; Conservative 51; Mismatches 101; Indels 11; Gaps 4;  
QY 37 CQOSLVK---QPMWNIFNFRILGSGQVEKSYPMQVSLKQKQKHICGGSTVSPQWVIT 93  
DB 848 CGIRMVNKKSEPAVGSRRFFSRISWRNSTVTGHPQVSLKSDHFCGSLIQEDRVVT 907  
QY 94 AAHCI--ANRNIVSTLNVTAGYDLSQTPGQTLTIETVIHPHFSTKPKMDYDIALLK 151  
DB 908 AAHCLDSLSEKQLKNITVTSGEYSLFQDKQEQNPVSKIIITHEPNSREYMSPDIALY 967  
QY 152 MAGAFQGHFVGPICLPDLPELREQFAGFICTTAGWRLTEGGVLSQVLOEVNLPILTWECC 211  
DB 968 LKHVKFGNAVQPICLPDSDDKVEPILCLSSGKGKISKTSYSNVLOEMELPIMDDRAC 1027  
QY 212 VAALLTLKRPISGKTFLCTGPDGGRDACQGDGSGSLMCRNKKGAWTLAGVTSWLGCGR 107  
DB 1028 NTVLKSMLNPLPLGRTMLCAGFPDWMGACQGDGSGPLVCRGGGIWILAGITSWVAGCAG 1087  
QY 272 G---WRNVKRSQDQSGPGIFTDISKVLWIHEHIOTG 305  
DB 1088 GSVPRVNNHVKA---SLGIFSKVSELMDFITQNLFTG 1121

## RESULT 9

ABG10218  
ID ABG10218 standard; Protein: 1576 AA.

XX AC ABG10218;

XX AC ABG10218;

XX AC ABG10218;

XX AC ABG10218;

XX AC ABG10218;

XX AC ABG10218;

XX AC ABG10218;

XX AC ABG10218;

XX AC ABG10218;

XX AC ABG10218;

XX AC ABG10218;

XX AC ABG10218;

XX AC ABG10218;

PA (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX N-PSDB; AAS74405.

XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity

XX Claim 20; SEQ ID No 40577; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,  
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
XX and gene mapping, and in recombinant production of (II). The  
XX polynucleotides are also used in diagnostics as expressed sequence tags  
XX for identifying expressed genes. (I) is useful in gene therapy techniques  
XX to restore normal activity of (II) or to treat disease states involving  
XX (II). (II) is useful for generating antibodies against it, detecting or  
XX quantitating a polypeptide in tissue, as molecular weight markers and as  
XX a food supplement. (II) and its binding partners are useful in medical  
XX imaging of sites expressing (II). (I) and (II) are useful for treating  
XX disorders involving aberrant protein expression or biological activity.  
XX The polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. ABG00010-ABC30377 represent novel human  
XX diagnostic amino acid sequences of the invention.  
XX Note: The sequence data for this patent did not appear in the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 1576 AA;

Query Match 34.7%; Score 567.5; DB 22; Length 1576;  
Best Local Similarity 41.2%; Pred. No. 3.3e-39;  
Matches 114; Conservative 51; Mismatches 101; Indels 11; Gaps 4;

QY 37 CQOSLVK---QPMWNIFNFRILGSGQVEKSYPMQVSLKQKQKHICGGSTVSPQWVIT 93

DB 848 CGIRMVNKKSEPAVGSRRFFSRISWRNSTVTGHPQVSLKSDHFCGSLIQEDRVVT 907

QY 94 AAHCI--ANRNIVSTLNVTAGYDLSQTPGQTLTIETVIHPHFSTKPKMDYDIALLK 151

DB 908 AAHCLDSLSEKQLKNITVTSGEYSLFQDKQEQNPVSKIIITHEPNSREYMSPDIALY 967

QY 152 MAGAFQGHFVGPICLPDLPELREQFAGFICTTAGWRLTEGGVLSQVLOEVNLPILTWECC 211

DB 968 LKHVKFGNAVQPICLPDSDDKVEPILCLSSGKGKISKTSYSNVLOEMELPIMDDRAC 1027

QY 212 VAALLTLKRPISGKTFLCTGPDGGRDACQGDGSGSLMCRNKKGAWTLAGVTSWLGCGR 107

DB 1028 NTVLKSMLNPLPLGRTMLCAGFPDWMGACQGDGSGPLVCRGGGIWILAGITSWVAGCAG 1087

QY 272 G---WRNVKRSQDQSGPGIFTDISKVLWIHEHIOTG 305

DB 1088 GSVPRVNNHVKA---SLGIFSKVSELMDFITQNLFTG 1121

## RESULT 10

ABG14588  
ID ABG14588 standard; Protein: 1576 AA.

XX AC ABG14588;

XX AC ABG14588;

XX AC ABG14588;

XX AC ABG14588;

XX AC ABG14588;

XX AC ABG14588;

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US08631.  
XX  
XX 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.  
XX  
XX (HYSE-) HYSEQ INC.  
PA Drmanac RT, Liu C, Tang YT;  
PI WPI; 2001-639362/73.  
DR N-PSDB; AAS78775.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
XX Claim 20; SEQ ID No 44947; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 1576 AA;  
Query Match 34.7%; Score 567.5; DB 22; Length 1576;  
Best Local Similarity 41.2%; Pred. No. 3.3e-39;  
Matches 114; Conservative 51; Mismatches 101; Indels 11; Gaps 4;  
QY 37 CGOSLVK---QPWNYFNIFSRILGSGQVEKGYPMQVSLKQKQKHICGSGIVSPQWVIT 93  
DB 848 CGIRMYNMKKEPAGVGSRRFSRISWRNSTVTGHPMQVSLKSDHHFCCGSLIQEDRVVT 907  
QY 94 AARCI--ANRNVSTLNTAGYDLSQTDPGQTLTIETVTHPHFSTKPKMDYDIALLK 151  
DB 908 AAHCLDSLSEKQLKNITVTSGEYSLFQDKQEQNIPVSKLIITHPEVNSREYMSPDIALLY 967  
QY 152 MAGAFQFHGVFICPLPELREOFEGAGICTTAGWGRLEGGVLSQVLOEVNLPILTWEBC 211  
DB 968 LKHVKFGNAVQICPLPDSDRRDVEPGILCSGCGKISKITSEYSNVLQEMELPIMDDRAC 1027  
QY 212 VAALLTLKRPISGKTFCLCTGFGDGDACQSGSLMCRNKGAWTLGAVTSWGLGCCR 271  
DB 1028 NTVLKSNNLPPILGRITMLCAGFPDWGMDACQSGGLVLCVRGGGIWILAGITTSWAGCAG 1087  
QY 272 G---WRNNVRKSDQSGPGIFTDLSKVLWSIHEHIQTG 305

DB 1088 GSPVVRNHHVKA---SLGIFSKVSELMDFITQNLFTG 1121  
RESULT 11  
ABG19887  
ID ABG19887 standard; Protein; 1576 AA.  
XX  
XX AC ABG19887;  
XX  
XX 18-FEB-2002 (first entry)  
DE Novel human diagnostic protein #19878.  
XX  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
XX Homo sapiens.  
XX  
XX WO200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US08631.  
PR 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.  
XX  
XX (HYSE-) HYSEQ INC.  
PA Drmanac RT, Liu C, Tang YT;  
PI WPI; 2001-639362/73.  
DR N-PSDB; AAS84074.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
XX Claim 20; SEQ ID No 50246; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 1576 AA;  
Query Match 34.7%; Score 567.5; DB 22; Length 1576;  
Best Local Similarity 41.2%; Pred. No. 3.3e-39;  
Matches 114; Conservative 51; Mismatches 101; Indels 11; Gaps 4;  
QY 37 CGOSLVK---QPWNYFNIFSRILGSGQVEKGYPMQVSLKQKQKHICGSGIVSPQWVIT 93  
DB 848 CGIRMYNMKKEPAGVGSRRFSRISWRNSTVTGHPMQVSLKSDHHFCCGSLIQEDRVVT 907









**THIS PAGE BLANK (USPTO)**

Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	688.5	42.1	1004	2	T30338	oviductin (EC 3.4.21.1)
2	626	38.3	1524	2	T30337	polyprotein - African swine fever virus
3	486	29.7	786	1	A47547	serine proteinase
4	466	28.5	638	1	KQHUP	plasma kallikrein
5	464	28.4	638	1	KQMSPL	plasma kallikrein
6	447.5	27.4	638	1	KQRTPL	plasma kallikrein
7	444	27.1	343	1	A57014	proctasin (EC 3.4.21.1)
8	442.5	27.0	417	1	S00845	hepsin (EC 3.4.21.1)
9	442.5	27.0	1035	1	A43090	enteropeptidase (EC 3.4.21.1)
10	438.5	26.8	416	1	S33777	hepsin (EC 3.4.21.1)
11	438.5	26.8	1034	1	A53663	enteropeptidase (EC 3.4.21.1)
12	433	26.5	593	2	S45281	coagulation factor
13	431.5	26.4	1019	1	A56318	enteropeptidase (EC 3.4.21.1)
14	421	25.7	855	2	JC7731	membrane-bound arginase
15	421	25.7	855	2	JC7775	membrane type-III integral protein
16	420.5	25.7	1047	2	A56617	masquerade precursor
17	419.5	25.6	237	1	TRCY1	trypsin (EC 3.4.21.1)
18	419	25.6	264	2	I38136	chymotrypsin-like coagulation factor
19	418	25.6	625	1	KFHU1	trypsin (EC 3.4.21.1)
20	415.5	25.4	275	2	S40007	trypsin (EC 3.4.21.1)
21	414.5	25.3	615	1	KFHU12	coagulation factor
22	413.5	25.3	263	2	A21195	chymotrypsin (EC 3.4.21.1)
23	409	25.0	263	2	A31299	chymotrypsin (EC 3.4.21.1)
24	409	25.0	275	2	S40005	trypsin (EC 3.4.21.1)
25	408.5	25.0	415	1	A34170	acrosin (EC 3.4.21.1)
26	408.5	25.0	603	2	S28941	coagulation factor
27	407.5	24.9	274	2	S35339	trypsin (EC 3.4.21.1)
28	406	24.8	254	1	TRWV3Y	trypsin-like protease
29	405	24.8	265	2	T15451	hypothetical protein





Db	412	LVSONHMCSSITIGROWILTAACHFDGIPYDPVWRVYGGIILNSETNKTPESSIKELII	471
Qy	134	HPHESTKKPMDYDIALKKNAGAFQGFHVGPICLPELRQFEAGFTCTTAGRGRLTEGV	193
Db	472	HQYTKMSEG-SYDIALIKLOTPLNYTEFORKPICPSKADNTIYNCWWTGMYTKERGE	530
Qy	194	LSOVLOBNPLILTWEBCVAALLTKRPISGKTFCLCTGPDGDRACODSGSILMCRNK	253
Db	531	TONILQKATPLVPNEECQK--KYRDYVITQMICAGKEGDIACKGDSGGPLVCKH-	586
Qy	254	KGAWTLAGVTSWGLGCGRGWRNNVRKSDQSGPIFTDISKVLSWTHEIQT	304
Db	587	SGRWQLVGTISWEGGAR-----KEQPGVYTKAEYIDWLEIKIOS	627

RESULT 7  
A57014  
prostatin (EC 3.4.21.-) precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 18-Jun-1999  
C:Accession: A57014; A54866  
R:Yu, J.X.; Chao, L.; Chao, J.  
J. Biol. Chem. 270, 13483-13489, 1995  
A:Title: Molecular cloning, tissue-specific expression, and cellular localization of hum  
A:Reference number: A57014; MUID:95286644; PMID:7768952  
A:Accession: A57014  
A:Status: translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-343 <RES>  
A:Cross-references: GB:I41351; NID:9862304; PIDN:AAC41759.1; PID:9862305  
A:Experimental source: prostate  
A:Note: parts of this sequence were determined by protein sequencing  
R:Yu, J.X.; Chao, L.; Chao, J.  
J. Biol. Chem. 269, 18843-18848, 1994  
A:Title: Prostatin is a novel human serine proteinase from seminal fluid. Purification,  
A:Reference number: A54866; MUID:94308140; PMID:8034638  
A:Accession: A54866  
A:Molecule type: protein  
A:Residues: 45-64 <YUA>  
C:Genetics:

[illegible]

A:Accession: A43090  
A:Status: nucleic acid sequence not shown; translated from GB/EMBL/DDDB  
A:Molecule type: mRNA  
A:Residues: 1-1035 <KIT>  
A:Cross-references: GB:U09859; NID:g746410; PIDN:AAB40026.1; PID:g746411  
A:Experimental source: small intestine  
A:LaValle, E.R.; Rehmentulla, A.; Racie, L.A.; DiBlasio, E.A.; Ferenz, C.; Grant, K.L.; J. Biol. Chem. 268, 23311-23317, 1993  
A:Title: Cloning and functional expression of a cDNA encoding the catalytic subunit of B  
A:Reference number: A48874; MUID:94043122; PMID:8226855  
A:Accession: A48874  
A:Molecule type: mRNA  
A:Residues: 801-1035 <LAV>  
A:Cross-references: GB:L19663; NID:g416131; PIDN:AAAL6035.1; PID:g416132  
A:Note: parts of this sequence, including the amino end of the mature protein, were conf  
R:Light, A.; Janska, H.  
J. Protein Chem. 10, 475-480, 1991  
A:Title: The amino-terminal sequence of the catalytic subunit of bovine enterokinase.  
A:Reference number: A61436; MUID:92189715; PMID:1799406  
A:Accession: A61436  
A:Molecule type: protein  
A:Residues: 801-807 'Y', 809-827 <LIG>  
C:Comment: The mechanism of association with the membrane of the intestinal brush border  
embrane attachment using a signal-anchor sequence.  
C:Comment: Conversion from membrane-bound to soluble forms may involve further processin  
C:Complex: mature enteropeptidase is variously reported to contain two (heavy and light)  
Lfide linked  
C:Function:  
A:Description: cleaves propeptide from trypsinogen to produce active trypsin  
A:Pathway: intestinal digestive hydrolase cascade  
C:Superfamily: enteropeptidase; C1r/C1s repeat homology; LDL receptor ligand-binding rep  
C:Keywords: glycoprotein; hydrolase; intestine; serine proteinase; transmembrane protein  
F:22-38/Domain: transmembrane #status predicted <TM>  
F:52-117/Product: enteropeptidase mini chain #status predicted <MC>  
F:118-800/Product: enteropeptidase heavy chain #status predicted <HC>  
F:199-236/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
F:358-520/Domain: MAM homology <MAM>  
F:542-647/Domain: C1r/C1s repeat homology <C1r>  
F:659-693/Domain: LDL receptor ligand-binding repeat homology <LDL2>  
F:694-799/Domain: scavenger receptor cysteine-rich domain homology #status atypical <SRC  
F:801-1030/Product: enteropeptidase light chain #status predicted <LC>  
F:801-1030/Domain: trypsin homology <TRY>  
F:116-147,170,194,223,263,264,404,456,486,519,550,646,698,722,741,762,864,903,965/Bindin  
F:788-912,826-842,926-993,957-972,983-1011/Disulfide bonds: #status predicted  
F:841,892,987/Active site: His, Asp, Ser #status predicted

Query Match 27.0%; Score 442.5; DB 1; Length 1035;  
Best Local Similarity 34.9%; Pred. No. 7.4e-32;  
Matches 96; Conservative 54; Mismatches 90; Indels 35; Gaps 9;

QY 36 SCGSLV--KVOPNNYIFSRILGSGVEKSYSPQVSLKQKQKHICGSGIVSPQWVIT 93  
Db 787 SCGKLLTQVEVP-----KIVGSDSREGAPWVWVAFDDQVCGASLSRDLWS 838  
QY 94 AAHCIAARNI-VSTLNTAGEYDLSQ-TDPGEQTLTIETVIIHPFSTKPKMDYDIALLK 151  
Db 839 AAHCYGRNNEPSKWKAVGLHNASNLTSQIETRLDQIVINPHYN-KRKNNDIAMMH 897  
QY 152 MAGAFQGHFVGPICLPDLREQFAGICTAGNGRTEGGVLSQVLOEYNLPILTWEBC 211  
Db 898 LEMKNVNTDVIQICLPDENQVFPFGRICSTAGNGALTYQGSTADVLQEADVPLLSNEK 957  
QY 212 VAAL---LTLKRPISKCTFLCTGFPDGRDACOGDSGSLMCRNKKGAWTLGCVTSWGL 267  
Db 958 QQQMPENYIT-----ENMVCAEYAGVDSCOGSDGSLMCO-ENNRLLAGVTSFGY 1009  
QY 268 GCGRGNRNVRKSDQSGPGIFTDISKVLSTHEHI 302  
Db 1010 QCALPNR-----PGYARVPRETIHQSL 1034

RESULT 10  
S33777

hepsin (EC 3.4.21.-) - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 18-Jun-1999  
C:Accession: S33777; S32013  
R:Farley, D.; Raymond, F.; Nick, H.  
Biochim. Biophys. Acta 1173, 350-352, 1993  
A:Title: Cloning and sequence analysis of rat hepsin, a cell surface serine proteina-  
A:Reference number: S33777; MUID:93305733; PMID:8318546  
A:Accession: S33777  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-416 <FAR>  
A:Cross-references: EMBL:X70900; NID:g57928; PIDN:CAA50256.1; PID:g57929  
C:Superfamily: hepsin; trypsin homology  
C:Keywords: hydrolase; liver; serine proteinase; transmembrane protein  
F:22-44/Domain: transmembrane #status predicted <TM>  
F:162-399/Domain: trypsin homology <TRY>  
F:187-203,290-359,321-337,348-380/Disulfide bonds: #status predicted  
F:202,256,352/Active site: His, Asp, Ser #status predicted

Query Match 26.8%; Score 438.5; DB 1; Length 416;  
Best Local Similarity 33.0%; Pred. No. 5.4e-32;  
Matches 102; Conservative 48; Mismatches 112; Indels 47; Gaps 11;

QY 11 KLILLGIVFFERKSAALSLPKAPSCGSLVKVQPNYFNIFSRILGSGQVEKSGYPWQ 70  
Db 129 RLDDVISVCDPCGRFLTAT---CQDGRRLPV-----DRIVGGQDSSLGRWPWQ 176  
QY 71 VLSKQRKHICGSGIVSPQWVITAAHCIAARNI-IVSTLNTAGEYDLSQTDPGEQTLTIE 129  
Db 177 VSLRYDCTHLGSGSLSGDWLTAACHCFPERNRLSRVLRVFAQA--VARTSPHAVQLGVQ 234  
QY 130 TVIIHPHF----STKPKMDYDIALLKAGAFQGHFVGPICLPDLREQFAGICTAG 184  
Db 235 AVIYHGGLPFRDPTDITENSNDIALVHLSSSLPTETIQVPCLPAAGQALVDGKVCVTG 294  
QY 185 WGLRTEGGLVSOVLQEVNLPILTWEQVVAALLTLKRP-----TSGKTFELCTGFPDGRD 238  
Db 295 WGNTOFGQAVVLOEARVPIISNEVC-----NSPDFYGNQIKPKMF-CAGYPEGID 346  
QY 239 ACQGDGSGSLMCRN----KGAWTLAGVTSWGLGCGRGNRNVRKSDQSGPGIFTDISKVL 295  
Db 347 ACQGDGSGHFCVEDRIGTSRWRLCGIVSWGTGAL-----ARK-----PGVYTKVIDER 396  
QY 296 SWIHEHTQT 304  
Db 397 EWIFQAIKT 405

RESULT 11  
A53663  
enteropeptidase (EC 3.4.21.9) precursor - pig  
N:Alternate names: enterokinase  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 07-Oct-1994 #sequence\_revision 09-Aug-1996 #text\_change 18-Jun-1999  
C:Accession: A53663  
R:Watsushima, M.; Ichinose, M.; Yahagi, N.; Kakei, N.; Tsukada, S.; Kuroka  
J. Biol. Chem. 269, 19976-19982, 1994  
A:Title: Structural characterization of porcine enteropeptidase.  
A:Reference number: A53663; MUID:94327548; PMID:8051081  
A:Accession: A53663  
A:Molecule type: mRNA  
A:Residues: 1-1034 <MAT>  
A:Cross-references: GB:D30799; NID:g505122; PIDN:BAA06459.1; PID:g505123  
A:Note: parts of this sequence, including the amino ends of three chains isolated fr  
C:Comment: The mechanism of association with the membrane of the intestinal brush bo  
ated below) or with amino-terminal myristoylation of the heavy chain.  
C:Complex: Mature enteropeptidase is variously reported to contain two (heavy and li  
ed by a disulfide bond. Possibly, conversion from membrane-bound to soluble forms in  
C:Function:  
A:Description: cleaves activation peptide from trypsinogen to produce active trypsin  
A:Pathway: intestinal digestive hydrolase cascade  
C:Superfamily: enteropeptidase; C1r/C1s repeat homology; LDL receptor ligand-binding

C:Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein; zymogen  
F:22-38/Domain: transmembrane #status predicted <TM>  
F:52-117/Product: enteropeptidase mini chain #status predicted <MCH>  
F:118-799/Product: enteropeptidase heavy chain #status predicted <HCH>  
F:199-236/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
F:357-519/Domain: MAM homology <MAM>  
F:541-646/Domain: C1r/C1s repeat homology <C1r>  
F:658-692/Domain: LDL receptor ligand-binding repeat homology <LDL2>  
F:693-798/Domain: scavenger receptor cysteine-rich domain homology #status atypical <SRC>  
F:800-1034/Product: enteropeptidase light chain #status predicted <LCH>  
F:800-1029/Domain: trypsin homology <TRY>  
F:116,147,170,194,283,343,350,403,455,485,518,549,645,697,701,721,740,761,804,863,902,96  
F:787-911,825-841,925-992,996-971,982-1010/Disulfide bonds: #status predicted  
F:840,891,986/Active site: His, Asp, Ser #status predicted

Query Match 26.8%; Score 438.5; DB 1; Length 1034;  
Best Local Similarity 30.3%; Pred. No. 1.7e-31;  
Matches 101; Conservative 62; Mismatches 93; Indels 77; Gaps 11;

QY 15 LLGI-----VFFERKSAALSLPKAP-----SCQSILV 42  
Db 733 LLGLGTGSSMFFSSGGGPFVKLTAPNGSLILTAEQCFEDSLILQCNHKSCGKKQV 792

QY 43 --KVQPNWYFNFSRLGSGQVEKSGYPQVSLKQKQKHICGGSIVSQWVITAHCIAN 100  
Db 793 AOEVSF-----KIVGNDREGAWPMVVALYNYQLCCGASLVSRDLVSAARCVYG 844

QY 101 RNI-----VSTLVNVTAGEYDLSQTDGQETLTETVTHPHFSTKPKMDYDIALLKMA 153  
Db 845 RNLEPSKWKAILGLHMTS-----NLTSQIVTRLIDEIVINPHYNNRR-KDSDIAMHLE 898

QY 154 GAFQGFHVGPICLPDLREQFAGICTAGWGRTEGGVLSQVLQEVNLPILTWEQVA 213  
Db 899 FKNVYTDYIQTLPEENQVFPFGICISAGWGIYQGSFADILQEAADVPLLSNEKQQ 958

QY 214 AL----LTLKRPISGKTFLCTGFPDGGDACOGDGGSLMCRNKKGAWTLAGVTSWGLGC 269  
Db 959 QMPEYNIT-----ENMWCAGEGGIDSCQDGGSLMCL-ENNRWLLAGVTSFGVQC 1010

QY 270 GRGWRNVKRSQDGSFGIFTDISKVLWISHEHI 302  
Db 1011 ALPNR-----PGVVARVPKFTWISQSL 1033

RESULT 12  
S45281  
coagulation factor XIIfa (EC 3.4.21.38) precursor - bovine (fragment)  
N:Alternate names: Hageman factor (activated)  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 10-Apr-1995 #sequence\_revision 22-Apr-1995 #text\_change 21-Jan-2000  
C:Accession: S45281; A61329  
R:Shibuya, Y.; Semba, U.; Okabe, H.; Kambara, T.; Yamamoto, T.  
Biochim. Biophys. Acta 1206, 63-70, 1994  
A:Title: Primary structure of bovine Hageman factor (blood coagulation factor XII): comp  
A:Reference number: S45281; MUID:94242782; PMID:8186251  
A:Accession: S45281  
A:Molecule type: mRNA  
A:Residues: 1-593 <SH1>  
A:Cross-references: GB:S70164  
A:Note: the authors translated the codon GAG for residue 23 as Val, GAG for residue 70 a  
s Pro, CTC for residue 203 as Phe, GTG for residue 247 as Leu, CCG for residue 286 as Gl  
Is, and ATC for residue 505 as Leu  
R:Fujikawa, K.; Walsh, K.A.; Davie, E.W.  
Biochemistry 16, 2270-2278, 1977  
A:Title: Isolation and characterization of bovine factor XII (Hageman factor).  
A:Reference number: A61329; MUID:77182112; PMID:861210  
A:Accession: A61329  
A:Molecule type: protein  
A:Residues: 10-16,'X',18-19;525-550 <FUJ>  
C:Superfamily: coagulation factor XII; EGF  
C:Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; monomer; plasma; s  
F:37-78/Domain: fibronectin type II repeat homology <IF2>  
F:88-120/Domain: EGF homology <EGF>

F:125-160/Domain: fibronectin type I repeat homology <FB1>  
F:207-287/Domain: kringle homology <KR>  
F:350-587/Domain: trypsin homology <TRY>  
F:541/Active site: Ser #status predicted

Query Match 26.5%; Score 433; DB 2; Length 593;  
Best Local Similarity 36.1%; Pred. No. 2.7e-31;  
Matches 101; Conservative 45; Mismatches 94; Indels 40; Gaps 12;

QY 37 CGOSLVKQVPWNWYFNFSRLGSGQVEKSGYPQVSLKQKQKHICGGSIVSQWVITAHAH 96  
Db 336 CQQRURK-----WLSLNKRVGGLVALGAHPYIAALYWDQ-HFCAGSLIAPCWVLTAAH 389

QY 97 CIANRNVSTLVNVTAGEYDLSQTDGQETLTETVTHPHFSTKPKMDY--DIALLKMA 154  
Db 390 CLQNPAPKELTVLGDQRHNSCEQCOTLAVRDYRLHEAFS--PITYQHDALVLRLOE 446

QY 155 AFQ--FGH---FVGPICILPE--LRQFEAGFTCTTAGWGRLEGGVLSQVLQEVNLPILT 207  
Db 447 SADGCAHPSPFPQVCLPSTAAARPAEAAVCEVAGMHQFEGGEYSFLQEAQVPLID 506

QY 208 WEECVAAALLTLKRPISGKTF-----LCTGFPDGGDACOGDGGSLMCRNK--KGAWTLA 261  
Db 507 PORCSAP-----DVHGAFTQGLCAGFLEGGTDACQDGGPLVCEDETPEPQLILRG 560

QY 262 VTSWGLGCGRGWRNVKRSQDGSFGIFTDISKVLWISHEH 301  
Db 561 IVSWSGCG---NRLK-----PGVYTDVANYLAWIREH 590

RESULT 13  
A56318  
enteropeptidase (EC 3.4.21.9) precursor - human  
N:Alternate names: enterokinase  
C:Species: Homo sapiens (man)  
C:Date: 19-May-1995 #sequence\_revision 09-Aug-1996 #text\_change 18-Jun-1999  
C:Accession: A56318; B43090  
R:Kitamoto, Y.; Veille, R.A.; Donis-Keller, H.; Sadler, J.E.  
Biochemistry 34, 4562-4568, 1995  
A:Title: cDNA sequence and chromosomal localization of human enterokinase, the proteo  
A:Reference number: A56318; MUID:95234679; PMID:7718557  
A:Accession: A56318  
A:Molecule type: mRNA  
A:Residues: 1-1019 <KIT>  
A:Cross-references: GB:U09860; NID:g746412; PIDN:AAC50138.1; PID:g746413  
R:Kitamoto, Y.; Yuan, X.; Wu, Q.; McCourt, D.W.; Sadler, J.E.  
Proc. Natl. Acad. Sci. U.S.A. 91, 7588-7592, 1994  
A:Title: Enterokinase, the initiator of intestinal digestion, is a mosaic protease co  
A:Reference number: A43090; MUID:94329561; PMID:8052624  
A:Accession: B43090  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 749-1019 <K12>  
A:Cross-references: GB:U09860  
A:Comment: The mechanism of association with the membrane of the intestinal brush bor  
otated below) or with amino-terminal myristoylation of the heavy chain.  
C:Genetics:  
A:Gene: GDB:PRSS7  
A:Cross-references: GDB:384083; OMIM:226200  
A:Map position: 21q21-21q21  
C:Complex: Mature enteropeptidase is variously reported to contain two (heavy and lig  
ed by a disulfide bond. Possibly, conversion from membrane-bound to soluble forms inv  
ducts.  
C:Function:  
A:Description: cleaves activation peptide from trypsinogen to produce active trypsin  
A:Pathway: intestinal digestive hydrolase cascade  
C:Superfamily: enteropeptidase; C1r/C1s repeat homology; LDL receptor ligand-binding  
C:Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein; zymoge  
F:1-784/Product: enteropeptidase heavy chain #status predicted <HCH>  
F:22-38/Domain: transmembrane #status predicted <TM>  
F:184-221/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
F:342-504/Domain: MAM homology <MAM>  
F:526-631/Domain: C1r/C1s repeat homology <C1r>



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 03:59:00 ; Search time 26 Seconds  
(without alignments)  
629.210 Million cell updates/sec

Title: US-09-735-713a-2  
Perfect score: 1636  
Sequence: 1 MSKMLISRNKLLLLGIVF.....IFTDISKVLWTHIQTGN 306

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 221153 seqs, 53462247 residues  
Total number of hits satisfying chosen parameters: 221153

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :		Published_Applications_AA:*	
1:	/cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*	2:	/cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*	4:	/cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*	6:	/cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*	8:	/cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
9:	/cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*	10:	/cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
11:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*	12:	/cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*	14:	/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1636	100.0	306	9 US-09-735-713a-2	Sequence 2, Appli
2	1618	98.9	302	9 US-09-735-713a-4	Sequence 4, Appli
3	1382	83.3	556	10 US-09-888-615-106	Sequence 106, App
4	898	54.9	164	9 US-09-735-713a-6	Sequence 6, Appli
5	489.5	29.9	802	9 US-09-978-295A-169	Sequence 169, App
6	489.5	29.9	802	9 US-09-978-697-169	Sequence 169, App
7	489.5	29.9	802	9 US-09-978-192A-169	Sequence 169, App
8	489.5	29.9	802	9 US-09-999-832A-169	Sequence 169, App
9	489.5	29.9	802	9 US-09-978-189-169	Sequence 169, App
10	489.5	29.9	802	9 US-09-978-608A-169	Sequence 169, App
11	489.5	29.9	802	9 US-09-978-191A-169	Sequence 169, App
12	489.5	29.9	802	9 US-09-978-403A-169	Sequence 169, App
13	489.5	29.9	802	9 US-09-978-564A-169	Sequence 169, App
14	489.5	29.9	802	9 US-09-978-585A-169	Sequence 169, App
15	489.5	29.9	802	9 US-10-017-081A-169	Sequence 169, App
16	489.5	29.9	802	10 US-09-888-615-113	Sequence 113, App
17	467	28.5	453	9 US-09-978-295A-69	Sequence 69, Appl
18	467	28.5	453	9 US-09-978-697-69	Sequence 69, Appl
19	467	28.5	453	9 US-09-978-192A-69	Sequence 69, Appl

20	467	28.5	453	9 US-09-999-832A-69	Sequence 69, Appl
21	467	28.5	453	9 US-09-978-189-69	Sequence 69, Appl
22	467	28.5	453	9 US-10-174-590-64	Sequence 64, Appl
23	467	28.5	453	9 US-10-176-758-64	Sequence 64, Appl
24	467	28.5	453	9 US-10-175-737-64	Sequence 64, Appl
25	467	28.5	453	9 US-10-173-706-64	Sequence 64, Appl
26	467	28.5	453	9 US-10-175-738-64	Sequence 64, Appl
27	467	28.5	453	9 US-10-175-752-64	Sequence 64, Appl
28	467	28.5	453	9 US-10-176-482-64	Sequence 64, Appl
29	467	28.5	453	9 US-10-176-757-64	Sequence 64, Appl
30	467	28.5	453	9 US-10-176-913-64	Sequence 64, Appl
31	467	28.5	453	9 US-10-180-552-64	Sequence 64, Appl
32	467	28.5	453	9 US-10-180-557-64	Sequence 64, Appl
33	467	28.5	453	9 US-10-173-700-64	Sequence 64, Appl
34	467	28.5	453	9 US-10-174-572-64	Sequence 64, Appl
35	467	28.5	453	9 US-10-174-579-64	Sequence 64, Appl
36	467	28.5	453	9 US-10-174-582-64	Sequence 64, Appl
37	467	28.5	453	9 US-10-174-588-64	Sequence 64, Appl
38	467	28.5	453	9 US-10-175-739-64	Sequence 64, Appl
39	467	28.5	453	9 US-10-175-740-64	Sequence 64, Appl
40	467	28.5	453	9 US-10-175-743-64	Sequence 64, Appl
41	467	28.5	453	9 US-10-176-488-64	Sequence 64, Appl
42	467	28.5	453	9 US-10-176-492-64	Sequence 64, Appl
43	467	28.5	453	9 US-10-176-747-64	Sequence 64, Appl
44	467	28.5	453	9 US-10-176-750-64	Sequence 64, Appl
45	467	28.5	453	9 US-10-176-985-64	Sequence 64, Appl

ALIGNMENTS

RESULT 1  
US-09-735-713a-2  
; Sequence 2, Application US/09735713A  
; Patent No. US20020165376A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Turner, C. Alexander Jr.  
; APPLICANT: Abuin, Alejandro  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: Polynucleotides Encoding the Same  
; FILE REFERENCE: LEX-0108-USA  
; CURRENT APPLICATION NUMBER: US/09/735,713A  
; CURRENT FILING DATE: 2000-12-12  
; PRIOR FILING DATE: 1999-12-22  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 306  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-735-713a-2

Query Match		100.0%;	Score 1636;	DB 9;	Length 306;
Best Local Similarity		100.0%;	Pred. No. 5.1e-149;		
Matches 306;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MSKMLISRNKLLLLGIVFFERKSAALS	PKAPSCGQSLVKVQPNWNYNIFSRILGGS	60	
DB	1	MSKMLISRNKLLLLGIVFFERKSAALS	PKAPSCGQSLVKVQPNWNYNIFSRILGGS	60	
QY	61	QVEKGSYPMWVSLKQKQKHCIGGSIVSPQWVITA	AHCIAANRNIVSTLNVTAAGEYDLSQTD	120	
DB	61	QVEKGSYPMWVSLKQKQKHCIGGSIVSPQWVITA	AHCIAANRNIVSTLNVTAAGEYDLSQTD	120	
QY	121	PGEQTLTIETVIIHPHFSTKKPKMDYDIAL	LKMAGAFQGHVFGPICLPRLREOFEAGFIC	180	
DB	121	PGEQTLTIETVIIHPHFSTKKPKMDYDIAL	LKMAGAFQGHVFGPICLPRLREOFEAGFIC	180	

```
Qy 181 TTAGGRLTEGGVLSQVLQEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGGRDAC 240
|||||
Db 181 TTAGGRLTEGGVLSQVLQEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGGRDAC 240
|||||
Qy 241 QDGGSLMCRNKKGAWTLAGVTSWGLCGGRWRNNVRKSDGSPGIFTDISKVLWSIHE 300
|||||
Db 241 QDGGSLMCRNKKGAWTLAGVTSWGLCGGRWRNNVRKSDGSPGIFTDISKVLWSIHE 300
|||||
Qy 301 HIQTGN 306
|||||
Db 301 HIQTGN 306
|||||

RESULT 2
US-09-735-713A-4
; Sequence 4, Application US/09735713A
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abuin, Alejandro
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020165376A1 Human Proteases and
; FILE REFERENCE: Lex-0108-USA
; CURRENT APPLICATION NUMBER: US/09735,713A
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,566
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-735-713A-4

Query Match 98.9%; Score 1618; DB 9; Length 302;
Best Local Similarity 100.0%; Pred. No. 2.7e-147; Indels 0; Gaps 0;
Matches 302; Conservative 0; Mismatches 0;

Qy 5 MLISRNKLLILGIVFFERKSAALSLPKAPSCGQSLVKVQPNWYFNIFSRILGGSQVEK 64
|||||
Db 1 MLISRNKLLILGIVFFERKSAALSLPKAPSCGQSLVKVQPNWYFNIFSRILGGSQVEK 60
|||||
Qy 65 GSYPMQVSLKQKQKHICGGSIVSPQWVITAHCITANRNIVSTLNTVTAAGEYDLSQTDGEEQ 124
|||||
Db 61 GSYPMQVSLKQKQKHICGGSIVSPQWVITAHCITANRNIVSTLNTVTAAGEYDLSQTDGEEQ 120
|||||
Qy 125 TLTETVTHPHFTKPKMDYDIALKMGAFQGFHFGVPGICLPDLREQFAGFICTTAG 184
|||||
Db 121 TLTETVTHPHFTKPKMDYDIALKMGAFQGFHFGVPGICLPDLREQFAGFICTTAG 180
|||||
Qy 195 WGRTEGGVLSQVLQEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGGRDACQDGS 244
|||||
Db 181 WGRTEGGVLSQVLQEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGGRDACQDGS 240
|||||
Qy 245 GGSMLCRNKKGAWTLAGVTSWGLCGGRWRNNVRKSDGSPGIFTDISKVLWSIHEIQT 304
|||||
Db 241 GGSMLCRNKKGAWTLAGVTSWGLCGGRWRNNVRKSDGSPGIFTDISKVLWSIHEIQT 300
|||||
Qy 305 GN 306
||
Db 301 GN 302

RESULT 3
US-09-888-615-106
; Sequence 106, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
```

```
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 106
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-106

Query Match 83.3%; Score 1362; DB 10; Length 556;
Best Local Similarity 85.7%; Pred. No. 2e-122; Indels 8; Gaps 2;
Matches 263; Conservative 8; Mismatches 28;

Qy 1 MSLKMLISRNKLLILGIVFFERKSAALSLPKAPSCGQSLVKVQPNWYFNIFSRILGGS 60
|||||
Db 1 MSLKMLISRNKLLILGIVFFERKSAALSLPKAPSCGQSLVKVQPNWYFNIFSRILGGS 60
|||||
Qy 61 QVEKGSYPWQVSLKQKQKHICGGSIVSPQWVITAHCITANRNIVSTLNTVTAAGEYDLSQTD 120
|||||
Db 61 QVEKGSYPWQVSLKQKQKHICGGSIVSPQWVITAHCITANRNIVSTLNTVTAAGEYDLSQTD 120
|||||
Qy 121 PGEOTLTETVTHPHFTKPKMDYDIALKMGAFQGFHFGVPGICLPDLREQFAGFIC 180
|||||
Db 121 PGEOTLTETVTHPHFTKPKMDYDIALKMGAFQGFHFGVPGICLPDLREQFAGFIC 180
|||||
Qy 181 TTAGWRLTEGGVLSQVLQEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGGRDAC 240
|||||
Db 181 TTAGWRLTEGGVLSQVLQEVNLPILTWEECVAAALLTLKRPISGKTFCTGFPDGGRDAC 240
|||||
Qy 241 QDGGSLMCRNKKGAWTLAGVTSWGLCGGRWRNNVRKSDGSP-----GIFTDISK 293
|||||
Db 241 QDGGSLMCRNKKGAWTLAGVTSWGLCGGRWRNNVRKSDGSP-----GIFTDISK 293
|||||
Qy 294 VLSWIHE 300
||
Db 300 CGTWCE 306

RESULT 4
US-09-735-713A-6
; Sequence 6, Application US/09735713A
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abuin, Alejandro
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020165376A1 Human Proteases and
; FILE REFERENCE: Lex-0108-USA
; CURRENT APPLICATION NUMBER: US/09/735,713A
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,566
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Homo sapiens
```

US-09-735-713A-6

Query Match 54.9%; Score 898; DB 9; Length 164;  
Best Local Similarity 100.0%; Pred. No. 1.2e-78;  
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 MDYDIALKMGAFQGFHFVGPICLPELREQFEAGFICTAGWGRLTGGVLSQVLQEVN 202  
|||||  
Db 1 MDYDIALKMGAFQGFHFVGPICLPELREQFEAGFICTAGWGRLTGGVLSQVLQEVN 60  
QY 203 LPILTWEECVAALLTKRPISGKTEFLCTGFPDGGRDACQDGGSGSLMCRNKKGANTLAGV 262  
|||||  
Db 61 LPILTWEECVAALLTKRPISGKTEFLCTGFPDGGRDACQDGGSGSLMCRNKKGANTLAGV 120  
QY 263 TSWGLGCGGRWNNVRKSDQSGPGIFTDISKVLNHIHQTN 306  
|||||  
Db 121 TSWGLGCGGRWNNVRKSDQSGPGIFTDISKVLNHIHQTN 164

RESULT 5

US-09-978-295A-169  
; Sequence 169, Application US/09978295A  
; Patent No. US20020156006A1

GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan  
; APPLICANT: Ferrara, Napoleon  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Kuo, Sophia S.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Shelton, David L.  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: P2630P1C11

; CURRENT APPLICATION NUMBER: US/09/978,295A

; CURRENT FILING DATE: 2001-10-15

; PRIOR APPLICATION NUMBER: 09/918585

; PRIOR FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/064249

; PRIOR FILING DATE: 1997-11-03

; PRIOR APPLICATION NUMBER: 60/065311

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/066364

; PRIOR FILING DATE: 1997-11-21

; PRIOR APPLICATION NUMBER: 60/077450

; PRIOR FILING DATE: 1998-03-10

; PRIOR APPLICATION NUMBER: 60/077632

; PRIOR FILING DATE: 1998-03-11

; PRIOR APPLICATION NUMBER: 60/077641

; PRIOR FILING DATE: 1998-03-11

; PRIOR APPLICATION NUMBER: 60/077649

; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077791  
; PRIOR FILING DATE: 1998-03-12  
; PRIOR APPLICATION NUMBER: 60/078004  
; PRIOR FILING DATE: 1998-03-13  
; PRIOR APPLICATION NUMBER: 60/078886  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078936  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078939  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079294  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: 60/079656  
; PRIOR FILING DATE: 1998-03-26  
; PRIOR APPLICATION NUMBER: 60/079664  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079689  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079663  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079728  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079786  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079920  
; PRIOR FILING DATE: 1998-03-30  
; PRIOR APPLICATION NUMBER: 60/079923  
; PRIOR FILING DATE: 1998-03-30  
; PRIOR APPLICATION NUMBER: 60/080105  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080107  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080165  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080194  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080327  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080328  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080333  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080334  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/081070  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081049  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081071  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081195  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081203  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081229  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081955  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081817  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081819  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081952  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081838  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/082568  
; PRIOR FILING DATE: 1998-04-21  
; PRIOR APPLICATION NUMBER: 60/082569  
; PRIOR FILING DATE: 1998-04-21







; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339

; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 29.9%; Score 489.5; DB 9; Length 802;
Best Local Similarity 39.3%; Pred. No. 1.4e-38;
Matches 99; Conservative 44; Mismatches 86; Indels 23; Gaps 5;

QY 54 SRILGGSQVEKGSYPQVSLKQKHICGSGSIVSPQWVITAAHRIANRNVSTL--NVTA 111
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 566 SRIVGGVSSGEWPMQASLYVRGHRICGALLADRWVITAAHCFQEDSMSTVLTWTVFL 625
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 112 GEYDLSQTDPDGEOTLTITETVTHPHFSTKKPMYDIALKMGAFQFHFVGICLPCLR 171
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 626 GKWQNSRWPGEVSFKVSRLLHP-YHEEDSHDYDVALQLDHPVYRVAARVPCLPARS 684
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 172 EQFEAGFICTTAGWGLTEGCVLSQVLQEVNLPILTWECEVAALLTLKRPISGKTPCLTCG 231
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 685 HFPEPLHCWITGWCALREGGPISNALQKVDYQLIPQDLCSEAYRYQVTP----RMLCAG 740
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 232 FPDGGRDACQDGSGLMCRNKKGAWTLAGVTSWGICGGRWNNVRKSDQGS---GIF 288
| : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 741 YRKGGKDACQDGSGLVCKALSGRWFAGLVSWGLGCGR-----PNYFGYV 787
| : | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 289 TDISKVLSWIE 300
| | | | | :
DB 788 TRITGVISWQQ 799
| | | | | :

RESULT 8
US-09-999-832A-169
; Sequence 169, Application US/09999832A
; Publication No. US20020192706A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann

APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2630P1G63  
CURRENT APPLICATION NUMBER: US/09/999,832A  
CURRENT FILING DATE: 2001-10-24  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334

PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084639  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640  
PRIOR FILING DATE: 1998-05-07





```

; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT:
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC22
; CURRENT APPLICATION NUMBER: US/09/978,608A
; CURRENT FILING DATE: 2001-10-16
; NUMBER OF SEQ ID NOS: 624
; Prior Application removed - See File Wrapper or Palm
; SEQ ID NO 169
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-798-608A-169

Query Match 29.9%; Score 489.5; DB 9; Length 802;
Best Local Similarity 39.3%; Pred. No. 1.4e-38;
Matches 99; Conservative 44; Mismatches 86; Indels 23; Gaps 5;

QY 54 SRILGSGVEKSGSPQVSLKQKQKHCIGGSIIVSPQWVITAAHCIAANRIVSTL--NVTA 111
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 566 SRIVGAVSSEGEWPQASLQVRGRHICGALIDRWVITAAHCFQEDSMASTVLTWFL 625
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 112 GEYDLSQTDPGQTLTIETVITHPHFSTKKPMDYDIALKMGAGAFQGFHFGVPICLPELR 171
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 626 GKWNQNSRMPGEVSEKSLLLHP-YHEEDSHDYDVALQLDHPVVRSAARVPCLPARS 684
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 172 EQFEAGFICTTAGWRLTEGGVLSQVLQVNLPILTWEECVAAALLTLKRPISGKTFCLCTG 231
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 685 HFEFGLHCWITGALREGGPISNALQKVDVQLIPQDLTSEAYRYQVTP----RMLCAG 740
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 232 FPDGGRDACQSGSGLMCRNKKKGAWTLGAVTWSLGGCGRGNRNVRKSDQSP---GIF 288
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 741 YRKGKDACQSGSGGLVCKALSGRFLAGLSVGLGCGR-----PNYFGVY 787
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
QY 289 TDISKVLSIHE 300
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 788 TRITGVISWIIQ 799
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 11
US-09-798-191A-169
; Sequence 169, Application US/09978191A
; Publication No. US20030050239A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT:
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC22
; CURRENT APPLICATION NUMBER: US/09/978,191A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
```

; Sequence 169, Application US/09978403A





7	PRIOR FILING DATE: 1998-04-28	
7	PRIOR APPLICATION NUMBER: 60/083392	
7	PRIOR FILING DATE: 1998-04-29	
7	PRIOR APPLICATION NUMBER: 60/083495	
7	PRIOR FILING DATE: 1998-04-29	
7	PRIOR APPLICATION NUMBER: 60/083496	
7	PRIOR FILING DATE: 1998-04-29	
7	PRIOR APPLICATION NUMBER: 60/083499	
7	PRIOR FILING DATE: 1998-04-29	
7	PRIOR APPLICATION NUMBER: 60/083545	
7	PRIOR FILING DATE: 1998-04-29	
7	PRIOR APPLICATION NUMBER: 60/083554	
7	PRIOR FILING DATE: 1998-04-29	
7	PRIOR APPLICATION NUMBER: 60/083558	
7	PRIOR FILING DATE: 1998-04-29	
7	PRIOR APPLICATION NUMBER: 60/083559	
7	PRIOR FILING DATE: 1998-04-29	
7	PRIOR APPLICATION NUMBER: 60/083500	
7	PRIOR FILING DATE: 1998-04-29	
7	PRIOR APPLICATION NUMBER: 60/083742	
7	PRIOR FILING DATE: 1998-04-30	
7	PRIOR APPLICATION NUMBER: 60/084366	
7	PRIOR FILING DATE: 1998-05-05	
7	PRIOR APPLICATION NUMBER: 60/084414	
7	PRIOR FILING DATE: 1998-05-06	
7	PRIOR APPLICATION NUMBER: 60/084441	
7	PRIOR FILING DATE: 1998-05-06	
7	PRIOR APPLICATION NUMBER: 60/084637	
7	PRIOR FILING DATE: 1998-05-07	
7	PRIOR APPLICATION NUMBER: 60/084639	
7	PRIOR FILING DATE: 1998-05-07	
7	PRIOR APPLICATION NUMBER: 60/084640	
7	PRIOR FILING DATE: 1998-05-07	
7	PRIOR APPLICATION NUMBER: 60/084640	
7	PRIOR FILING DATE: 1998-05-07	
7	PRIOR APPLICATION NUMBER: 60/084598	
7	PRIOR FILING DATE: 1998-05-07	
7	PRIOR APPLICATION NUMBER: 60/084600	
7	PRIOR FILING DATE: 1998-05-07	
7	PRIOR APPLICATION NUMBER: 60/084627	
7	PRIOR FILING DATE: 1998-05-07	
7	PRIOR APPLICATION NUMBER: 60/084643	
7	PRIOR FILING DATE: 1998-05-07	
7	PRIOR APPLICATION NUMBER: 60/085339	
7	PRIOR FILING DATE: 1998-05-13	
7	PRIOR APPLICATION NUMBER: 60/085338	
7	PRIOR FILING DATE: 1998-05-13	
7	PRIOR APPLICATION NUMBER: 60/085323	
7	PRIOR FILING DATE: 1998-05-13	
7	PRIOR APPLICATION NUMBER: 60/085582	
7	PRIOR FILING DATE: 1998-05-15	
7	PRIOR APPLICATION NUMBER: 60/085700	
7	PRIOR FILING DATE: 1998-05-15	
7	PRIOR APPLICATION NUMBER: 60/085689	
7	PRIOR FILING DATE: 1998-05-15	
7	PRIOR APPLICATION NUMBER: 60/085573	
7	PRIOR FILING DATE: 1998-05-15	
7	PRIOR APPLICATION NUMBER: 60/085704	
7	PRIOR FILING DATE: 1998-05-15	
7	PRIOR APPLICATION NUMBER: 60/085697	

Seq. ID	Seq. Name	Seq. Length	Seq. Type	Seq. Source	Seq. Accession	Seq. Date	Seq. Status	Seq. Notes
1	Seq. 1	100	Protein	Human	NC_000001	2000	Active	
2	Seq. 2	100	Protein	Human	NC_000002	2000	Active	
3	Seq. 3	100	Protein	Human	NC_000003	2000	Active	
4	Seq. 4	100	Protein	Human	NC_000004	2000	Active	
5	Seq. 5	100	Protein	Human	NC_000005	2000	Active	
6	Seq. 6	100	Protein	Human	NC_000006	2000	Active	
7	Seq. 7	100	Protein	Human	NC_000007	2000	Active	
8	Seq. 8	100	Protein	Human	NC_000008	2000	Active	
9	Seq. 9	100	Protein	Human	NC_000009	2000	Active	
10	Seq. 10	100	Protein	Human	NC_000010	2000	Active	
11	Seq. 11	100	Protein	Human	NC_000011	2000	Active	
12	Seq. 12	100	Protein	Human	NC_000012	2000	Active	
13	Seq. 13	100	Protein	Human	NC_000013	2000	Active	
14	Seq. 14	100	Protein	Human	NC_000014	2000	Active	
15	Seq. 15	100	Protein	Human	NC_000015	2000	Active	
16	Seq. 16	100	Protein	Human	NC_000016	2000	Active	
17	Seq. 17	100	Protein	Human	NC_000017	2000	Active	
18	Seq. 18	100	Protein	Human	NC_000018	2000	Active	
19	Seq. 19	100	Protein	Human	NC_000019	2000	Active	
20	Seq. 20	100	Protein	Human	NC_000020	2000	Active	
21	Seq. 21	100	Protein	Human	NC_000021	2000	Active	
22	Seq. 22	100	Protein	Human	NC_000022	2000	Active	
23	Seq. 23	100	Protein	Human	NC_000023	2000	Active	
24	Seq. 24	100	Protein	Human	NC_000024	2000	Active	
25	Seq. 25	100	Protein	Human	NC_000025	2000	Active	
26	Seq. 26	100	Protein	Human	NC_000026	2000	Active	
27	Seq. 27	100	Protein	Human	NC_000027	2000	Active	
28	Seq. 28	100	Protein	Human	NC_000028	2000	Active	
29	Seq. 29	100	Protein	Human	NC_000029	2000	Active	
30	Seq. 30	100	Protein	Human	NC_000030	2000	Active	
31	Seq. 31	100	Protein	Human	NC_000031	2000	Active	
32	Seq. 32	100	Protein	Human	NC_000032	2000	Active	
33	Seq. 33	100	Protein	Human	NC_000033	2000	Active	
34	Seq. 34	100	Protein	Human	NC_000034	2000	Active	
35	Seq. 35	100	Protein	Human	NC_000035	2000	Active	
36	Seq. 36	100	Protein	Human	NC_000036	2000	Active	
37	Seq. 37	100	Protein	Human	NC_000037	2000	Active	
38	Seq. 38	100	Protein	Human	NC_000038	2000	Active	
39	Seq. 39	100	Protein	Human	NC_000039	2000	Active	
40	Seq. 40	100	Protein	Human	NC_000040	2000	Active	
41	Seq. 41	100	Protein	Human	NC_000041	2000	Active	
42	Seq. 42	100	Protein	Human	NC_000042	2000	Active	
43	Seq. 43	100	Protein	Human	NC_000043	2000	Active	
44	Seq. 44	100	Protein	Human	NC_000044	2000	Active	
45	Seq. 45	100	Protein	Human	NC_000045	2000	Active	
46	Seq. 46	100	Protein	Human	NC_000046	2000	Active	
47	Seq. 47	100	Protein	Human	NC_000			

[illegible]

```
Db 626 GKWNQSRWPGVEYFKVSRLLLLHP-YHEEDSHDYDVALLQDHPVVRSAARVPVCLPARS 684
QY 172 EQEAGFICTAGWRTGGVLSQVLQEVNLPILTWEECVAALLTLKRPISGKTFELCTG 231
Db 685 HFPEGLHCWITGALREGGPIISNALQKVDVQLIPODLCLSEAYRYQVTP-----RMLCAG 740
QY 232 FPDGRRDACOGDGGSLMCRNKKGANTLAGVTSWGLCGGRGWRNNVRKSDQSGSP---GIF 288
Db 741 YRKGKDACOGDGGPLVCKALSGRWFLAGLVSWGLCGGR-----PNYFGVY 787
QY 289 TDISKVLSWIHE 300
Db 788 TRITGVISWIQQ 799

RESULT 14
US-09-978-585A-169
; Sequence 169, Application US/09978585A
; Publication No. US20030049633A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC15
; CURRENT APPLICATION NUMBER: US/09/978,585A
; CURRENT FILING DATE: 2001-10-16
; NUMBER OF SEQ ID NOS: 624
; Prior Application removed - See File Wrapper or Palm
; SEQ ID NO 169
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-978-585A-169

Query Match 29.9%; Score 489.5; DB 9; Length 802;
Best Local Similarity 39.3%; Pred. No. 1.4e-38;
Matches 99; Conservative 44; Mismatches 86; Indels 23; Gaps 5;

QY 54 SRLGGSQVEKGSYPWQVSLKQKHICGGSIVSPQWVITAACHIANRNIVSTL--NVTA 111
Db 566 SRVGGAVSSEGEWPMQASLQVRGRHICGALLADRWVITAACHCFQEDSMASVLTWTVFL 625
QY 112 GEYDLSQTDGCEQTLTIETVIIHPHFSTKKPMDYDIALKMGAFQFGHFVGPICLPFLR 171
Db 626 GKWNQSRWPGVEYFKVSRLLLLHP-YHEEDSHDYDVALLQDHPVVRSAARVPVCLPARS 684
QY 172 EQEAGFICTAGWRTGGVLSQVLQEVNLPILTWEECVAALLTLKRPISGKTFELCTG 231
```

```
Db 685 HFPEGLHCWITGALREGGPIISNALQKVDVQLIPODLCLSEAYRYQVTP-----RMLCAG 740
QY 232 FPDGRRDACOGDGGSLMCRNKKGANTLAGVTSWGLCGGRGWRNNVRKSDQSGSP---GIF 288
Db 741 YRKGKDACOGDGGPLVCKALSGRWFLAGLVSWGLCGGR-----PNYFGVY 787
QY 289 TDISKVLSWIHE 300
Db 788 TRITGVISWIQQ 799

RESULT 15
US-10-017-081A-169
; Sequence 169, Application US/10017081A
; Publication No. US20030049684A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC69
; CURRENT APPLICATION NUMBER: US/10/017,081A
; CURRENT FILING DATE: 2002-04-30
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 624
; SEQ ID NO 169
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-081A-169

Query Match 29.9%; Score 489.5; DB 9; Length 802;
Best Local Similarity 39.3%; Pred. No. 1.4e-38;
Matches 99; Conservative 44; Mismatches 86; Indels 23; Gaps 5;

QY 54 SRLGGSQVEKGSYPWQVSLKQKHICGGSIVSPQWVITAACHIANRNIVSTL--NVTA 111
Db 566 SRVGGAVSSEGEWPMQASLQVRGRHICGALLADRWVITAACHCFQEDSMASVLTWTVFL 625
QY 112 GEYDLSQTDGCEQTLTIETVIIHPHFSTKKPMDYDIALKMGAFQFGHFVGPICLPFLR 171
Db 626 GKWNQSRWPGVEYFKVSRLLLLHP-YHEEDSHDYDVALLQDHPVVRSAARVPVCLPARS 684
QY 172 EQEAGFICTAGWRTGGVLSQVLQEVNLPILTWEECVAALLTLKRPISGKTFELCTG 231
Db 685 HFPEGLHCWITGALREGGPIISNALQKVDVQLIPODLCLSEAYRYQVTP-----RMLCAG 740
```

QY

232

FPDGRDACQDSGGSLMCRNKKGAWTLAGVTSWGLGCGRGWRNNVVKSDQSP---

GIF

288

Db

741

YRKGRKDACQDSGGPLVCCKALSGRWF LAGLVSWGLGCGR-----

PNYFGVY

787

QY

289

TDISKVLSWHE

300

Db

788

TRITGVISWIOQ

799

Search completed: March 25, 2003, 04:15:59

Job time : 30 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 03:57:25 ; Search time 22 Seconds  
(without alignments)  
409.246 Million cell updates/sec

Title: US-09-735-713A-2

Perfect score: 1636

Sequence: 1 MSKMLISRNKLLILLGIVF.....IFTDISKVLWIHIQTGN 306

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	%			DB	ID	Description
	Score	Match	Length			
1	467	28.5	454	4	US-09-518-046-2	Sequence 2, Appli
2	460	28.1	248	4	US-08-944-483-63	Sequence 63, Appli
3	447.5	27.4	638	2	US-08-681-151-3	Sequence 3, Appli
4	446.5	27.3	455	4	US-09-261-416-2	Sequence 2, Appli
5	444.5	27.2	328	4	US-09-386-642-11	Sequence 11, Appli
6	442.5	27.0	798	1	US-08-200-900A-2	Sequence 2, Appli
7	442.5	27.0	798	5	PCR-US94-00616-2	Sequence 2, Appli
8	440.5	26.9	416	2	US-09-000-846-2	Sequence 12, Appli
9	438	26.8	319	4	US-09-386-642-12	Sequence 3, Appli
10	437.5	26.7	256	2	US-09-027-337-3	Sequence 3, Appli
11	437.5	26.7	256	4	US-09-644-600-3	Sequence 3, Appli
12	432.5	26.4	255	4	US-08-944-483-67	Sequence 67, Appli
13	432.5	26.4	400	4	US-09-004-731-30	Sequence 30, Appli
14	432.5	26.4	400	4	US-09-004-731-33	Sequence 33, Appli
15	432.5	26.4	400	4	US-08-749-699-30	Sequence 30, Appli
16	432.5	26.4	400	4	US-08-749-699-33	Sequence 33, Appli
17	432.5	26.4	400	4	US-09-004-729-30	Sequence 30, Appli
18	432.5	26.4	400	4	US-09-004-729-33	Sequence 33, Appli
19	430.5	26.3	387	4	US-09-032-215-8	Sequence 8, Appli
20	430.5	26.3	387	4	US-09-032-215-13	Sequence 13, Appli
21	430	26.3	299	4	US-08-944-483-66	Sequence 66, Appli
22	430	26.3	315	4	US-09-386-653A-9	Sequence 9, Appli
23	429	26.2	230	4	US-09-386-653A-7	Sequence 7, Appli
24	426	26.0	242	4	US-09-004-731-36	Sequence 36, Appli
25	426	26.0	242	4	US-08-749-699-36	Sequence 36, Appli
26	426	26.0	242	4	US-09-004-729-36	Sequence 36, Appli
27	425.5	26.0	235	3	US-08-807-151-3	Sequence 3, Appli

28	425.5	26.0	235	4	US-09-478-957-3	Sequence 3, Appli
29	421.5	25.8	317	4	US-09-386-629-7	Sequence 7, Appli
30	419.5	25.6	235	4	US-08-944-483-65	Sequence 65, Appli
31	419.5	25.6	237	4	US-08-163-919A-3	Sequence 3, Appli
32	419.5	25.6	237	5	PCR-US94-14073-3	Sequence 3, Appli
33	418	25.6	327	4	US-09-386-629-8	Sequence 8, Appli
34	414	25.3	435	4	US-09-008-271A-6	Sequence 2, Appli
35	413	25.2	423	4	US-09-656-002-2	Sequence 6, Appli
36	408	24.9	238	4	US-08-944-483-64	Sequence 64, Appli
37	408	24.9	283	3	US-08-807-151-1	Sequence 1, Appli
38	408	24.9	283	4	US-09-478-957-1	Sequence 1, Appli
39	407	24.9	492	4	US-09-342-749-2	Sequence 2, Appli
40	407	24.9	492	4	US-09-691-840-2	Sequence 2, Appli
41	406	24.8	254	4	US-09-578-303-5	Sequence 5, Appli
42	405	24.8	231	2	US-09-027-337-6	Sequence 6, Appli
43	405	24.8	231	4	US-09-644-600-6	Sequence 6, Appli
44	403	24.6	314	4	US-09-008-271A-3	Sequence 3, Appli
45	402	24.6	267	2	US-09-016-366A-23	Sequence 23, Appli

ALIGNMENTS

RESULT 1  
US-09-518-046-2  
; Sequence 2, Application US/09518046  
; Patent No. 6294663  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed  
; TITLE OF INVENTION: in Ovarian Carcinoma and Uses Thereof  
; FILE REFERENCE: D6192CIP  
; CURRENT APPLICATION NUMBER: US/09/518,046  
; CURRENT FILING DATE: 2000-03-02  
; EARLIER APPLICATION NUMBER: 09/261,416  
; EARLIER FILING DATE: 1999-03-03  
; NUMBER OF SEQ ID NOS: 153  
; SEQ ID NO 2  
; LENGTH: 454  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: complete amino acid sequence of TADG-12  
; OTHER INFORMATION: protein  
US-09-518-046-2

Query Match	28.5%	Score 467;	DB 4;	Length 454;
Best Local Similarity	32.1%	Pred. No. 4.5e-42;		
Matches	99;	Conservative 61;	Mismatches 114;	Indels 34; Gaps 9;
QY	1	MSLKMLISRNKLILLGIVFFERG-KSAALSPLKAPSCGSLVKVQPNVFNFSRLGG	59	
	:	:	:	:
	:	:	:	:
DB	171	VSIDHLLPDKVTALHSHSVYREGCASHGVVTLQCTACGHRH-----GYSSRIVGG	221	
	:	:	:	:
	:	:	:	:
QY	60	SQVEKSGYPQVSLKQKQKHICGGSIVSPQWVITAAHCIAANRNIVSLNVTAGEYDLSQT	119	
	:	:	:	:
	:	:	:	:
DB	222	NMSLLSQWQASLQFGYHLCGGSVITPLWITAAHCIVTDLYLPKSWTQVGLVSLD-	280	
	:	:	:	:
	:	:	:	:
QY	120	DPGEQTLTTETVITHFSTKKPMDYDIALKMAQAFQFGFVGPICLPRLREQEAGFI	179	
	:	:	:	:
	:	:	:	:
DB	281	NPAPSHL-VEKIVYHSYKPKR-LGNDIALMKLAGLPTFNEMIQPVCLPSEENFPDGK	338	
	:	:	:	:
	:	:	:	:
QY	180	CTTAGWRLTEGGVLSVQLEVNLPILTWECEVAALLTLKRPISG---KTFICTGPPDG	235	
	:	:	:	:
	:	:	:	:
DB	339	CWTSGWGATEDGGDASPLNHAAPLISNKC-----NHRDVYGGIISPSMLCAGLYTG	392	
	:	:	:	:
	:	:	:	:
QY	236	GRDACGDSGSLMCRNKKGAWTLAGVTSNGLCGRGRNVRNKRKSDGSGCIFTDISKVL	295	
	:	:	:	:
	:	:	:	:
DB	393	GVDSQCQSDGSLPVCQBRR-LWKLVGATSFSGIC-----AEVNKPGVTRVTSFL	441	
	:	:	:	:
	:	:	:	:
QY	296	SWIHEHIQ 303		
	:	:	:	:

```
Db 442 DWHQEOME 449

RESULT 2
US-08-944-483-63
; Sequence 63, Application US/08944483
; Patent No. 6232456
; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.
; APPLICANT: STEWART, KENT D.
; APPLICANT: STROUPE, STEVEN D.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
; TITLE OF INVENTION: OF THE PROSTATE
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ FOR Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944.483
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6183 US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6232456e
US-08-944-483-63

Query Match 28.1%; Score 460; DB 4; Length 248;
Best Local Similarity 35.4%; Pred. No. 1.1e-41;
Matches 90; Conservative 50; Mismatches 96; Indels 18; Gaps 5;

Qy 56 ILGGSQVEKGYPMQVSLKOR---QKHICGGSGIVSPQWVITAHCIAANRNIVSTLNVTAG 112
Db 1 IVGGTSSNGEPMQVSLKORLTAQRHLCCGSLGHQWLVTAACHFDGLPLQDVWRIYSG 60
Qy 113 EYDLSQTDPGQTLTIEVILPHFSTKKPMDYDIALKMGAFQFGHFVGPICLPPELRE 172
Db 61 ILNLSIDTKDTPFSGIKKIIHQNTKVSQEG-NHDIALIKLOAPLNYTEFQKPICLPSKGD 119
Qy 173 QFAGFICTTAGGRLTEGGVLSQVLOEVNPLITWECVAALLTKRPISGKTFPLCTGPF 232
Db 120 TSTIYTNWVTGWGFSKEGEIQNLQKVNIPLVNTNECQKRYDYK---ITQRMVCAGY 176
Qy 233 PDGGRDACQSGSGSLMCRNKKGAWTLAGVTSWGLGCGRGRNVRNVRKSDQSGSPGIFTDIS 292

Db 177 KEGGKDACGDSGGPLVCKH-NGMWRVLGITSWGGCAR-----REQPGVYTKVA 225
Qy 293 KVLWIHEHIQTGN 306
Db 226 EYMDWILEKTQSSD 239

RESULT 3
US-08-681-151-3
; Sequence 3, Application US/08681151
; Patent No. 5869637
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Bandman, Olga
; APPLICANT: Braxton, Scott Michael
; APPLICANT: GOLL, Surya
; TITLE OF INVENTION: A NOVEL HUMAN KALLIKREIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/681.151
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0074US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 638 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 205011
US-08-681-151-3

Query Match 27.4%; Score 447.5; DB 2; Length 638;
Best Local Similarity 33.3%; Pred. No. 9.4e-40;
Matches 97; Conservative 49; Mismatches 118; Indels 27; Gaps 7;

Qy 18 IVFERGKSA-ALSLPKAPSCGQSLVKVQPNWNYENIFSRILGGSQVEKGYPMQVSLKOR 76
Db 360 ITYEAQSSGYSRLCKVVESSDCTTKIN-----ARIVGGTSSNGEPMQVSLQVK 411
Qy 77 ---QKHICGGSGIVSPQWVITAHCIAANRNIVSTLNVTAGYDLSDTDPGEQTLTETVII 133
Db 412 LVSNQHMCGGSIIGQWILTAAHCFDGPYPDVWRIYGGILNLSEITNKTFFSSIKELII 471
Qy 134 HPHSTKPKMDYDIALKMGAFQFGHFVGPICLPPELREQEAGFICTTAGRLETCGV 193
Db 472 HQYKMSQEG-SYDIALIKLOPLNYTEFQKPICLPSKADTNTIYTNWVTGWYTKERGE 530
```



```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 798 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-200-900A-2

Query Match          27.0%; Score 442.5; DB 1; Length 798;
Best Local Similarity 34.9%; Pred. No. 4.5e-39;
Matches 96; Conservative 54; Mismatches 90; Indels 35; Gaps 9;

QY 36 SCGSLV--KVQPNYNIIFSRILGGSGVEKGYSPWVSLKQKHKICGGSIYSPQWVIT 93
Db 550 SCGRKLVTEVSP-----KIVGSDSREGAPWVVALYFDQVCGASLSRDLVLS 601
QY 94 AAHCIAIRNI-VSTLNTAGYDLSQ-TDPEQTLTIETVLIHHPFSTKKPMDYDIALLK 151
Db 602 AAHCYGRNMEPSKWKAVLGJHMASNLTSPOIETRLIDQIVINPHYN-KRRKNDIAMMH 660
QY 152 MAGAFQFCHFGVPICLPELREQFEAGFTCTTAGMGRITGEGVLSQLQEVNLPILTWEEC 211
Db 661 LEMKVNTDYIQICLPEENQVPPGRICSTAGWALYQGSTADVLQEADVPLLSNEKC 720
QY 212 VAAL----LTLKRPISGKTFCTGFPDGGDRDACQDGGSLMCRNKKGAWTLAGVTSWGL 267
Db 721 QQQMPEYNIT-----ENMVACGYEAGGVDSGCGSGPLMCQ-ENNRWLLAGVTSFGY 772
QY 268 GCGRGWRNNVRKSDQSGPGIFTDISKVLWISHEHI 302
Db 773 QCALPNR-----PGVYARVPRTETWISQSL 797

RESULT 7
PCT-US94-00616-2
; Sequence 2, Application PC/TUS94000616
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: CLONING OF ENTEROKINASE AND METHOD OF USE
; NUMBER OF SEQUENCES: 33
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/00616
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 798 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-00616-2

Query Match          27.0%; Score 442.5; DB 5; Length 798;
Best Local Similarity 34.9%; Pred. No. 4.5e-39;
Matches 96; Conservative 54; Mismatches 90; Indels 35; Gaps 9;

QY 36 SCGSLV--KVQPNYNIIFSRILGGSGVEKGYSPWVSLKQKHKICGGSIYSPQWVIT 93
Db 550 SCGRKLVTEVSP-----KIVGSDSREGAPWVVALYFDQVCGASLSRDLVLS 601
QY 94 AAHCIAIRNI-VSTLNTAGYDLSQ-TDPEQTLTIETVLIHHPFSTKKPMDYDIALLK 151
Db 602 AAHCYGRNMEPSKWKAVLGJHMASNLTSPOIETRLIDQIVINPHYN-KRRKNDIAMMH 660
QY 152 MAGAFQFCHFGVPICLPELREQFEAGFTCTTAGMGRITGEGVLSQLQEVNLPILTWEEC 211
Db 661 LEMKVNTDYIQICLPEENQVPPGRICSTAGWALYQGSTADVLQEADVPLLSNEKC 720

; SEQUENCE CHARACTERISTICS:
; LENGTH: 798 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-000-846-2

Query Match          26.9%; Score 440.5; DB 2; Length 416;
Best Local Similarity 33.0%; Pred. No. 2.9e-39;
Matches 102; Conservative 48; Mismatches 112; Indels 47; Gaps 11;

QY 11 KLILLGIVFERGKSAALSAPRSCGQSLVKVQPNYNIIFSRILGGSGVEKGYSPWQ 70
Db 129 RLLDVISVDCPRGRFLTAT---CQDGRRLKPV-----DRIVGGQDSSLGRWPWQ 176
QY 71 VSLKQKHKICGGSIYSPQWVITAAHCIAIRNI-TVSTLNTAGYDLSQTDPPGQTLTIE 129
Db 177 VSURYDTHLGGSLSGDWLTAAHCFPERNRVLRSRWFAGA--VARTSPHAVQLGVQ 234
QY 130 TVIHPHF-----STKKPMDYDIALLKAGAFQFHFVGPICLPPELREQFEAGFTCTAG 184
Db 235 AVIVHGYGLFPROPTIDENSNDIALVHLSLPLTEVIQVCLPAAGQALVDGKVCITVG 294
QY 185 WGRITEGCVLSQVLQEVNLPILTWEECAVALLTLKRP-----ISGKTFCTGFPDGGDR 238
Db 295 WGNTOFYGOAMVLEQARVPIISNEVC-----NSPDFYGNQIKPKMF-CAGYPEGGID 346
```





Db 199 ITAAHCVAT-TPNSNLKVRLEWDYRDHDERLNHEEYAIATERKEVHPSYS---PTDFRNDV 254  
QY 148 ALLKMGAFQFGHFGVPTCLPELREQFAGFICTTAGWRLTEG-GVLSQLQEVNLPIL 206  
Db 255 ALVKLDRTVIFKQHLPLVCLPH-KOMKLAGKMATVAGWGRTRHGOSTVPAVLQEVNLPIL 313  
QY 207 TWECEVAALLTL-KRPSGKTLCTGFPDGGDRDACQSGGSLMCRNKKGAWTLAGVTSW 265  
Db 314 PNERCQRFRAAGRRETHDVFCLCAGYKGGDRSCQDGGGLIMQ-IEGRRTLGLVSW 372  
QY 266 GLGCGRGRNRRNKKSDQSGPGIFTDISKLSWI 298  
Db 373 GIGCGR-----EHLPGVYTNQKFPW 395  
RESULT 14  
US-09-004-731-33  
; Sequence 33, Application US/09004731  
; Patent No. 6177258  
; GENERAL INFORMATION:  
; APPLICANT: Wu Hunter, Shirley  
; APPLICANT: Stiegler, Gary  
; APPLICANT: Gaines, Patrick J.  
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheridan Ross P.C.  
; STREET: 1700 Lincoln Street, Suite 3500  
; CITY: Denver  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80203  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/004,731  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/749,699  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Connell, Gary J.  
; REGISTRATION NUMBER: 32,020  
; REFERENCE/DOCKET NUMBER: 2618-25-C3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 863-9700  
; TELEFAX: (303) 863-0223  
; INFORMATION FOR SEQ ID NO: 33:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 400 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-004-731-33  
Query Match 26.4%; Score 432.5; DB 4; Length 400;  
Best Local Similarity 37.0%; Pred. No. 2e-38;  
Matches 101; Conservative 46; Mismatches 91; Indels 35; Gaps 11;  
QY 36 SCGQSLVKQVPWNRYNIFSRILGGQVEKSGYSPWQVSLKQ----RKHICGGSIQSPQW 91  
Db 148 TCGE-----LYTRSNRIVGGHSTGFGSHPWQAALIKSGFLSKKLCGGLVSDRW 198  
QY 92 ITAAHCIAANRNVSTLNTVAGYDLSQDTP--GEOTLTIETVTHPHFTKKPMY--DI 147  
Db 199 ITAAHCVAT-TPNSNLKVRLEWDYRDHDERLNHEEYAIATERKEVHPSYS---PTDFRNDV 254  
QY 148 ALLKMGAFQFGHFGVPTCLPELREQFAGFICTTAGWRLTEG-GVLSQLQEVNLPIL 206  
Db 255 ALVKLDRTVIFKQHLPLVCLPH-KOMKLAGKMATVAGWGRTRHGOSTVPAVLQEVNLPIL 313  
QY 207 TWECEVAALLTL-KRPSGKTLCTGFPDGGDRDACQSGGSLMCRNKKGAWTLAGVTSW 265  
Db 314 PNERCQRFRAAGRRETHDVFCLCAGYKGGDRSCQDGGGLIMQ-IEGRRTLGLVSW 372

Db 255 ALVKLDRTVIFKQHLPLVCLPH-KOMKLAGKMATVAGWGRTRHGOSTVPAVLQEVNLPIL 313  
QY 207 TWECEVAALLTL-KRPSGKTLCTGFPDGGDRDACQSGGSLMCRNKKGAWTLAGVTSW 265  
Db 314 PNERCQRFRAAGRRETHDVFCLCAGYKGGDRSCQDGGGLIMQ-IEGRRTLGLVSW 372  
QY 266 GLGCGRGRNRRNKKSDQSGPGIFTDISKLSWI 298  
Db 373 GIGCGR-----EHLPGVYTNQKFPW 395  
RESULT 15  
US-08-749-699-30  
; Sequence 30, Application US/08749699  
; Patent No. 6210920  
; GENERAL INFORMATION:  
; APPLICANT: Wu Hunter, Shirley  
; APPLICANT: Stiegler, Gary  
; APPLICANT: Gaines, Patrick J.  
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sheridan Ross P.C.  
; STREET: 1700 Lincoln Street, Suite 3500  
; CITY: Denver  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80203  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/749,699  
; FILING DATE:  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Connell, Gary J.  
; REGISTRATION NUMBER: 32,020  
; REFERENCE/DOCKET NUMBER: 2618-25-C3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 863-9700  
; TELEFAX: (303) 863-0223  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 400 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-749-699-30  
Query Match 26.4%; Score 432.5; DB 4; Length 400;  
Best Local Similarity 37.0%; Pred. No. 2e-38;  
Matches 101; Conservative 46; Mismatches 91; Indels 35; Gaps 11;  
QY 36 SCGQSLVKQVPWNRYNIFSRILGGQVEKSGYSPWQVSLKQ----RKHICGGSIQSPQW 91  
Db 148 TCGE-----LYTRSNRIVGGHSTGFGSHPWQAALIKSGFLSKKLCGGLVSDRW 198  
QY 92 ITAAHCIAANRNVSTLNTVAGYDLSQDTP--GEOTLTIETVTHPHFTKKPMY--DI 147  
Db 199 ITAAHCVAT-TPNSNLKVRLEWDYRDHDERLNHEEYAIATERKEVHPSYS---PTDFRNDV 254  
QY 148 ALLKMGAFQFGHFGVPTCLPELREQFAGFICTTAGWRLTEG-GVLSQLQEVNLPIL 206  
Db 255 ALVKLDRTVIFKQHLPLVCLPH-KOMKLAGKMATVAGWGRTRHGOSTVPAVLQEVNLPIL 313  
QY 207 TWECEVAALLTL-KRPSGKTLCTGFPDGGDRDACQSGGSLMCRNKKGAWTLAGVTSW 265  
Db 314 PNERCQRFRAAGRRETHDVFCLCAGYKGGDRSCQDGGGLIMQ-IEGRRTLGLVSW 372

Qy 266 GLCGRGWRNNVRKSDQSPGIFTDISKVLWSI 298  
|:|||| : ||:|:| : ||  
Db 373 GIGGR-----EHLPGVVTNIQKFIPI 395

Search completed: March 25, 2003, 04:03:36  
Job time : 28 secs



```
Db 654 GTCTCATCATCACTCTTGCAGATGACAAAGGTGACTGCAATTACACCACCTCAGTATAT 713
QY 21 PheGluArgGly---LysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGln 39
Db 714 GTGAGGAGGAGTGCCTCTGCCACGTGGTTACCTTGAGTGCACACGCTGTGTGCAT 773
QY 40 SerLeuValLysValGlnProTrpAsnTrpPheAsnIlePheSerArgIleLeuGlyGly 59
Db 774 AGAAGG-----GGCTACAGCTCACGCATCGTGGGTGA 806
QY 60 SerGlnValGluLysGlySerTrpProTrpGlnValSerLeuLysGlnArgGlnLysHis 79
Db 807 AACATGCTCTTGCCTCGCAGTGCCTCGCAGCCAGCCAGCTTCAGTTCAGGGGCTACCAC 866
QY 80 IleCysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAla 99
Db 867 CTGTGGGGGGCTCTGTCATCAGCCCTGTGGATCATCTGCTGCACACTGTGTTTAT 926
QY 100 AsnArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTrpAspLeuSerGlnThr 119
Db 927 GACTTGATCTCCCAAGTCAAGCATCCAGGTGGGTCTAGTTTCCCTGTGGAC--- 983
QY 120 AspProGlyGluGlnThrIleGluThrValIleIleHisProHisPheSerThr 139
Db 984 AATCCAGCCCCATCCCACTTG---GTGGAGAAAGATTGTTTACCACAGCAAGTACAAGCCA 1040
QY 140 LysLysProMetAspTrpAspIleAlaLeuLysMetAlaGlyAlaPheGlnPheGly 159
Db 1041 AAGAGG---CTGGCAATGACATCGCCCTTATGAGCTGGCGGGCCACTCAGTTCAAT 1097
QY 160 HisPheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIle 179
Db 1098 GAAATGATCCAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1157
QY 180 CysThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGln 199
Db 1158 TGCTGAGCTCAGGATGGGGCCACAGAGGATGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1217
QY 200 GluValAsnLeuProIleLeuThrTrpGluGlu-CysValAlaAlaLeuLeuThrLeuLys 219
Db 1218 CACGGCGCGTCTCTTGTATTTCCAAAGAGATCTGC-----AACCA 1259
QY 219 sArgProIleSerGly-----LysThrPheLeuCysThrGlyPheProAspG1 235
Db 1260 CAGGGACGTGACGTGGCAGTCACTCTCCCTCCATGCTGCTGCTGCTGCTGCTGCTGCTG 1318
QY 235 yGlyArg---AspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLys 254
Db 1319 GTGGCTTGGAAACAGCTGCCAGGGGGACAGCGGGGGGGCCCTGGTGTGTCAAGAGAGGAG 1378

; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 1
; LENGTH: 2416
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: 144..1511
; OTHER INFORMATION: CDS
US-09-261-416-1

Alignment Scores:
Pred. No.: 1,73e-40 Length: 2416
Score: 446.50 Matches: 101
Percent Similarity: 52.28% Conservatives: 61
Best Local Similarity: 32.58% Mismatches: 112
Query Match: 27.29% Indels: 37
DB: 4 Gaps: 10

US-09-735-713A-2 (1-306) x US-09-261-416-1 (1-2416)
QY 1 MetSerLeuLysMetLeuIleSerArgAsnLysLeuIleLeuLeuLeuGlyIleValPhe 20
Db 654 GTGTCATCATCACTCTTGCAGATGACAAAGGTGACTGCAATTACACCACCTCAGTATAT 713
QY 21 PheGluArgGly---LysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGln 39
Db 714 GTGAGGAGGAGTGCCTCTGCCACGTGGTTACCTTGAGTGCACACGCTGTGTGCAT 773
QY 40 SerLeuValLysValGlnProTrpAsnTrpPheAsnIlePheSerArgIleLeuGlyGly 59
Db 774 AGAAGG-----GGCTACAGCTCACGCATCGTGGGTGA 806
QY 60 SerGlnValGluLysGlySerTrpProTrpGlnValSerLeuLysGlnArgGlnLysHis 79
Db 807 AACATGCTCTTGCCTCGCAGTGCCTCGCAGCCAGCCAGCTTCAGTTCAGGGGCTACCAC 866
QY 80 IleCysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAla 99
Db 867 CTGTGGGGGGCTCTGTCATCAGCCCTGTGGATCATCTGCTGCACACTGTGTTTAT 926
QY 100 AsnArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTrpAspLeuSerGlnThr 119
Db 927 GACTTGATCTCCCAAGTCAAGCATCCAGGTGGGTCTAGTTTCCCTGTGGAC--- 983
QY 120 AspProGlyGluGlnThrIleGluThrValIleIleHisProHisPheSerThr 139
Db 984 AATCCAGCCCCATCCCACTTG---GTGGAGAAAGATTGTTTACCACAGCAAGTACAAGCCA 1040
QY 140 LysLysProMetAspTrpAspIleAlaLeuLysMetAlaGlyAlaPheGlnPheGly 159
Db 1041 AAGAGG---CTGGCAATGACATCGCCCTTATGAGCTGGCGGGCCACTCAGTTCAAT 1097
QY 160 HisPheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIle 179
Db 1098 GAAATGATCCAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1157
QY 180 CysThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGln 199
Db 1158 TGCTGAGCTCAGGATGGGGCCACAGAGGATGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1217
QY 200 GluValAsnLeuProIleLeuThrTrpGluGlu-CysValAlaAlaLeuLeuThrLeuLys 219
Db 1218 CACGGCGCGTCTCTTGTATTTCCAAAGAGATCTGC-----AACCCAC 1259
QY 220 ArgProIleSerGly-----LysThrPheLeuCysThrGlyPheProAspGly 235
Db 1260 AGGACGTGATCGGTGGGATCATCTCCCTCCATGCTGCTGCGGGGCTACCTGACGGGT 1319
QY 236 GlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGly 255
Db 1320 GCGGTGAACAGCTGCCAGGGGACAGCGGGGGCCCTGGTGTGTCAAGAGAGGAGG--- 1376
QY 256 AlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsn 275
Db 1377 CTGTGGAAGTTAGTGGGAGCGACCAAGCTTGGCATCGGCTGC----- 1418
QY 276 AsnValArgLysSerAspGlnGlySerProGlyIlePheThrAspIleSerLysValLeu 295
Db 1419 -----GCAGAGTGAAACAGAGCTGGGGGTGTACACCCGTGCACCTCTTCCTCTG 1466

QY 296 SerTrpIleHisGluHisGln 303
Db 1467 GACTGGATCCAGGACAGATGGAG 1490

RESULT 2
US-09-261-416-1
; Sequence 1, Application US/09261416A
; Patent No. 6291663
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: TAGD-12: A No. 6291663el Transmembrane Serine Protease
; FILE OF INVENTION: Overexpressed in Ovarian Carcinoma
; FILE REFERENCE: D6192
; CURRENT APPLICATION NUMBER: US/09/261,416A
```

```
QY 254 sGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpAr 274
Db 1379 G-----CTGTGAAGTTAGTGGGACGACAGCTTTGGCATCGGCTGC----- 1421
QY 274 gAsnAsnValArgLysSerAspGlnGlySerProGlyIlePheThrAspLysSerLysVa 294
Db 1422 -----GCAGACGTGAACAAGCTGGGTGGTGTACACCGGTGTACACCTCCCTT 1465
QY 294 lLeuSerTrpIleHisGluHisIleGln 303
Db 1466 CTGGACTGGATCCACGACGATGGAG 1493
RESULT 3
US-09-386-642-7
; Sequence 7, Application US/09386642
; Patent No. 6420157
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew
; APPLICANT: QL, Jensen
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: Zymogen Activation System
; FILE REFERENCE: ORT-1028
; CURRENT APPLICATION NUMBER: US/09/386,642
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1169
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene
; OTHER INFORMATION: with homo sapien serine protease catalytic domain
US-09-386-642-7
Alignment Scores:
Pred. No.: 9,63e-41 Length: 1169
Score: 444.50 Matches: 100
Percent Similarity: 49.03% Conservative: 52
Best Local Similarity: 32.26% Mismatches: 119
Query Match: 27.17% Indels: 39
DB: 4 Gaps: 8
US-09-735-713A-2 (1-306) x US-09-386-642-7 (1-1169)
QY 3 LeuLysMetLeuIleSerArgAsnLysLeuIleLeuLeuLeuGlyIleValPhePheGlu 22
Db 46 CTGCTCTGCTGTGTGGTGTCAAAATCTACTCTTGTGCCAGGGTGTGGTCTCCGACTAC 105
QY 23 ArgGly-----LysSerAlaAlaLeuSerLeuProLysAlaProSerCysGly 38
Db 106 AAGGACGACGACGACGTCGGGCGCGCTCTGTGCGCCCTC----- 147
QY 39 GlnSerLeuValLysValGlnProTrpAsnTrpPheAsnIlePheSerArgIleLeuGly 58
Db 148 -----TTTGATGATGATGACACAGATCGTTGGG 174
QY 59 GlySerGlnValGluLysGlySerTrpTrpGlnValSerLeuLysGlnArgGlnLys 78
Db 175 GGCTATGCTCTAGAGCCGGGTGAGTGGCCCTGCAGGTACAGTACCTATGAAGCGCTC 234
QY 79 HisIleCysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIle 98
Db 235 CANGTGTGGTGGCTCTCTGCTGTGACAGTGGGTGCTGACGTCTGCTCTACTCTTC 294
QY 99 AlaAsnArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTrpAspLeuSerGln 118
Db 295 CCCAGCGAGCACCACACAGGAAGCTATGAGTCAAGCTGGGGGCCACCACTAGACTCC 354
QY 119 ThrAspProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSer 138
Db 355 TACTCCGAGGACGCCAAGGTCAGCACCCCTGAAGGACATCATCCCCACCCACGACTAC--- 411
```

```
QY 139 ThrLysLysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPhe 158
Db 412 CTCAGGAGGGCTCCCGAGGCGGACATTGCACTCTCCCACTCAGCAGCCCATCACCTTC 471
QY 159 GlyHisPheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPhe 178
Db 472 TCCCGCTACATCGGGCGCCATCTCCCTCCCTCGAGCCCAACGCCCTCTTCCCAAGGGCTC 531
QY 179 IleCysThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeu-----SerGln 196
Db 532 CACTGCACCTCTCACTGGCTGGGTGATGTGCCCCCTCAGTGAGCCCTCTCGACCCCAAG 591
QY 197 ValLeuGlnGluValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeu 216
Db 592 CCACTGCAGCAACTCGAGGTGCCTCTCATCATCGTGCAGACGCTGTAACCTGCCTGACAAC 651
QY 217 ThrLeuLysArgPro-----IleSerGlyLysThrPheLeuCysThrGlyPhe 232
Db 652 ATCGAGCCCAAGCTCGAGAGCGCACCTTTGTCCAGAGGACATGTTGTGTGTGGCTAT 711
QY 233 ProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArgAsn 252
Db 712 GTGGAGGGGGCAAGGACGCTGCCAGGTGACTCTGGGGGGCCACTCTCTCTGC---CCT 768
QY 253 LysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArgGly 272
Db 769 GTGGAGGGTCTCTGGTACCTGACGGGCATTTGTGAGCTGGGAGATGCTGTGGGGCC--- 825
QY 273 TrpArgAsnValArgLysSerAspGlnGlySerProGlyIlePheThrAspIleSer 292
Db 826 ---CGCAACAGG-----CCTGGTGTGTACTACTCTGGCCCTC 858
QY 293 LysValLeuSerTrpIleHisGluHisIle 302
Db 859 AGCTATGCTCTCTGGATCCAAAGCAAGGTG 888
RESULT 4
US-09-510-738A-188
; Sequence 188, Application US/09510738A
; Patent No. 6268165
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP-A
; CURRENT APPLICATION NUMBER: US/09/510,738A
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 188
; LENGTH: 1783
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: full length cDNA of hepsin
US-09-510-738A-188
Alignment Scores:
Pred. No.: 1.83e-40 Length: 1783
Score: 444.50 Matches: 103
Percent Similarity: 49.51% Conservative: 50
Best Local Similarity: 33.33% Mismatches: 120
Query Match: 27.17% Indels: 36
DB: 4 Gaps: 9
US-09-735-713A-2 (1-306) x US-09-510-738A-188 (1-1783)
QY 16 LeuGlyIleValPhePheGluArgGlyLysSerAla----- 27
Db 587 GTGGGGCTTCTTGTGTGGACGAGGGGAGGCTGCCACACCCAGAGGCTGTGGAGGT 646
QY 28 -----AlaLeuSerLeuProLysAlaProSerCysGlyGlnSerLeuValLysValGln 45
```



QY 276 nvalArgLysSerAspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSe 296  
Db 1984 -----CCAGGGCTCTACACCAAGTCACTGACTTCCGGGA 2018  
QY 296 rTrpIleHisGluHisIleGlnThr 304  
Db 2019 GTGATCTTCCAGGCCATAAAGACT 2043

## RESULT 6

US-08-200-900A-1  
; Sequence 1, Application US/08200900A  
; Patent No. 5665566  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: CLONING OF ENTEROKINASE AND METHOD OF USE  
; NUMBER OF SEQUENCES: 38  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genetics Institute, Inc. - Legal Affairs  
; STREET: 87 CambridgePark Drive  
; CITY: Cambridge  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02140  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/200,900A  
; FILING DATE: 23-FEB-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meinert, Maureen C.  
; REGISTRATION NUMBER: 31,544  
; REFERENCE/DOCKET NUMBER: GI 5201-FWC  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 876-1170 x8574  
; TELEFAX: (617) 876-5851  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2581 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-200-900A-1

Alignment Scores:  
Pred. No.: 5,42e-40 Length: 2581  
Score: 442.50 Matches: 96  
Percent Similarity: 54.55% Conservative: 54  
Best Local Similarity: 34.91% Mismatches: 90  
Query Match: 27.05% Indels: 35  
DB: 5 Gaps: 9

US-09-735-713a-2 (1-306) x US-08-200-900A-1 (1-2581)

QY 36 SerCysGlyGlnSerLeuVal-----LysValGlnProTrpAsnTyrPheAsnIlePhe 53  
Db 1649 TCATGTGGGAAAAAAGTGGTCAAGAGTTAGCCCG-----1687  
QY 54 SerArgIleLeuGlyGlySerGlnValGluLysSerTyrProTrpGlnValSerLeu 73  
Db 1688 ---AAGATTGTGGAGGAAGTACTCCAGAGAGGAGCCCTGGGCTGTGGTGTCTGT 1744  
QY 74 LysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpValIleThr 93  
Db 1745 TATTTCACGATCAACAGGTCTCGGGAGCTCTCTGTGTGACAGGGGATTCGCTGTGTG 1804  
QY 94 AlaAlaHisCysIleAlaAsnArgAsnIle---ValSerThrLeuAsnValThrAlaGly 112  
Db 1805 GCCGCCCACTGCGGTACGGGAGAGAAATATGAGCCGCTTAAGTGGAAAGCAGTCTAGGC 1864

QY 113 GluTyrAspLeuSerGln---ThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131  
Db 1865 CTGCATATGGCATCAAAATCTGACTCTCTCAGATAGAAACTAGTGTGATTGACCAAAAT 1924  
QY 132 IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLys 151  
Db 1925 GTCATAAACCCACACTACAAT---AAACGGAGAAAGAAACAAATGACATTCGCCATGTCAT 1981  
QY 152 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171  
Db 1982 CTTGAAATGAAAGTGAACATACACAGATTATATACAGCTATTTCTTTACCAGAGAAAT 2041  
QY 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly 191  
Db 2042 CAAGTTTTCCTCCAGGAAGAAATTTGTCTATTCTGGTGGGGGGGCACTTATATATCAA 2101  
QY 192 GlyValLeuSerGlnValLeuGlnValAsnLeuProIleLeuThrTrpGluGluCys 211  
Db 2102 GGTTCCTACTGCAGAGCTACTGCAAGAGCTGACGTTCCCTTCTATCAATGAGAAATGT 2161  
QY 212 ValAlaAlaLeu-----LeuThrLeuLysArgProIleSerGlyLysThrPhe 227  
Db 2162 CAACACAGATGCCAGAAATATAACATTACG-----GAAATATG 2200  
QY 228 LeuCysThrGlyPheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySer 247  
Db 2201 GTGTGTCCAGCTATGAAGCAGGAGGGTAGATTCTTGTACAGGGGGATTCAGGGGGACCA 2260  
QY 248 LeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeu 267  
Db 2261 CTCATGTGCCAA---GAAACACACAGATGGCTCTGGTGGCGTGACGTCATTTGGATAT 2317  
QY 268 GlyCysGlyArgGlyTrpArgAsnValArgLysSerAspGlnGlySerProGlyIle 287  
Db 2318 CAATGTGCCTGCTTAATCGC-----CCAGGGGTG 2347  
QY 288 PheThrAspIleSerLysValLeuSerTrpIleHisGluHisIle 302  
Db 2348 TATGCCCGGGTCCCAAGGTTCCACAGATGGGATACAAAGTTTTCYA 2392

## RESULT 7

PCT-US94-00616-1  
; Sequence 1, Application PC/TUS9400616  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: CLONING OF ENTEROKINASE AND METHOD OF USE  
; NUMBER OF SEQUENCES: 33  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/00616  
; FILING DATE:  
; CLASSIFICATION:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2581 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
PCT-US94-00616-1

Alignment Scores:  
Pred. No.: 5,42e-40 Length: 2581  
Score: 442.50 Matches: 96  
Percent Similarity: 54.55% Conservative: 54  
Best Local Similarity: 34.91% Mismatches: 90  
Query Match: 27.05% Indels: 35  
DB: 5 Gaps: 9

US-09-735-713A-2 (1-306) x PCT-US94-00616-1 (1-2581)

```
QY 36 SerCysGlyGlnSerLeuVal-----LysValGlnProTrpAsnTyrPheAsnIlePhe 53
Db 1649 TCATGTGGGAAACACTGTGGTACTCAAGAAATTAGCCG-----1687

QY 54 SerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu 73
Db 1688 ---AGATTGTCGAGGAAGTACTCCAGAGAAGGAGCCCTGGGCTGGTGTCTG 1744

QY 74 LysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpValIleThr 93
Db 1745 TATTTCGACGATCAACAGGTCTCGGAGCTTCTCTGTTGAGCAGGATGGCTGGTGTG 1804

QY 94 AlaAlaHisCysIleAlaAsnArgAsnIle---ValSerThrLeuAsnValThrAlaGly 112
Db 1805 GCCGCCCACTGCTGACGGAGAAATATGAGCCGCTCTAAGTGGAAAGCAGTGTAGGC 1864

QY 113 GluTyrAspLeuSerGln---ThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131
Db 1865 CTGCATATGGCATCAATCTGACTTCTCTCAGATAGAACCTAGTGTGATTGACCAAAAT 1924

QY 132 IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLys 151
Db 1925 GTCATAAACCCACACTACAAT---AAACGAGAAAGAACAAATGACATGCCATGATGCAT 1981

QY 152 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171
Db 1982 CTTGAAATGAAGTGAACACTACAGATTATATACAGCTATTGTTTACCAGAGAAAT 2041

QY 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly 191
Db 2042 CAAATTTTCCCCCAGGAAGAAATTTGTTCTATTGCTGGCTGGGGGCACTTATATCA 2101

QY 192 GlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTrpGluGlyCys 211
Db 2102 GGTCTCTACTGACAGCTACTGCAAGAGCTGACGTCCCTTCTCTCAATGAGAAATGT 2161

QY 212 ValAlaAlaLeu-----LeuThrLeuLysArgProIleSerGlyLysThrPhe 227
Db 2162 CAACAACAGATGCCAGAAATATAACATTAGC-----GAAAAATATG 2200

QY 228 LeuCysThrGlyPheProAspGlyArgAspAlaCysGlnGlyAspSerGlyGlySer 247
Db 2201 GTGTGTGAGGCTATGAAACAGAGGGGTAGATTCTTGTACGGGGGATTCAGGCGGACCA 2260

QY 248 LeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeu 267
Db 2261 CTCATGTGCCAA---GAAAAACAACAGATGGCTCTGGCTGGCTGACGTCAATTTGGATAT 2317

QY 268 GlyCysGlyArgGlyTrpArgAsnAsnValArgLysSerAspGlnGlySerProGlyIle 287
Db 2318 CAATGTGCACTGCCTTAATCGC-----CCAGGGGTG 2347

QY 288 PheThrAspIleSerLysValLeuSerTrpIleHisGluHisIle 302
Db 2348 TATGCCCCGGTCCCCAAGGTTTCACAGATGGATACAAAGTTTCTA 2392
```

## RESULT 8

```
US-09-000-846-1
; Sequence 1, Application US/09000846
; Patent No. 5981830
; GENERAL INFORMATION:
; APPLICANT: WU, QINGYU
; APPLICANT: SADLER, JASPER
; TITLE OF INVENTION: KNOCKOUT MICE AND THEIR PROGENY WITH
; NUMBER OF INVENTION: 8
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MILLER, WHITE, ZELANO & BRANTIGAN, P. C.
; STREET: 2200 CLARENDON BLVD, SUITE 1400
; CITY: ARLINGTON
```

```
STATE: VA
COUNTRY: US
ZIP: 22201
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/000,846
FILING DATE: 30-DEC-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/866,058
FILING DATE: 30-MAY-1997
BEST LOCAL SIMILARITY: 35.29%
ATTORNEY/AGENT INFORMATION:
NAME: LEBOVITZ, RICHARD M.
REGISTRATION NUMBER: 37,067
REFERENCE/DOCKET NUMBER: BERLX 65P1
TELEPHONE: 703-243-6333
TELEFAX: 703-243-6410
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1605 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: 35..1282
US-09-000-846-1

Alignment Scores:
Pred. No.: 4,42e-40 Length: 1605
Score: 440.50 Matches: 102
Percent Similarity: 49.48% Conservative: 41
Best Local Similarity: 35.29% Mismatches: 102
Query Match: 26.93% Indels: 45
DB: 2 Gaps: 10

US-09-735-713A-2 (1-306) x US-09-000-846-1 (1-1605)

QY 31 LeuProLysAlaProSerCysGlyGlnSerLeuValLysValGlnProTrpAsnTyrPhe 50
Db 471 CTGCCACCTGCCAA-GACTGTGGCGCAGAGAGCTGCCGGTG-----511

QY 51 AsnIlePheSerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGln 70
Db 512 -----GACCGCATTTGGGGGGGCCAGGACAGTCTCGGAGAGTGGCCGTGGCAG 562

QY 71 ValSerLeuLysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrp 90
Db 563 GTCAGCCTCGGTTATGATGGGACCCACCTCTGTGGGGGTCCCTGCTCTGGGACTGG 622

QY 91 VallIleThrAlaAlaHisCysIleAlaAsnArgAsn---IleValSerThrLeuAsnVal 109
Db 623 GTGCTGACTGTGCACATTGCTTTCCAGAGGGGAACCGGGTCTCTCGGTGGCGAGTA 682

QY 110 ThrAlaGlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGlu 129
Db 683 TTTGCTGTGTGCT-----GTAGCCCGGACCTCACCCATGTGTCAACTGGGGGTTCAG 736

QY 130 ThrValIleIleHisProHisPhe-----SerThrLysLysProMetAsp 144
Db 737 GCTGTGATCTATCATGGGGGCTACCTTCCCTTCGAGACCCCTACTATCGACGAAACAGC 796

QY 145 TyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHisPheValGlyPro 164
```

[illegible]

```

RESULT 9
US-09-386-642-8
; Sequence 8, Application US/09386642
; Patent No. 6420157
; GENERAL INFORMATION:
; APPLICANT: Darrow, Andrew
; APPLICANT: QI, Jensen
; APPLICANT: Andrade-Gordon, Patricia
; TITLE OF INVENTION: Zymogen Activation System
; FILE REFERENCE: ORT-1028
; CURRENT APPLICATION NUMBER: US/09/386,642
; CURRENT FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 60
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1142
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene
; OTHER INFORMATION: with homo sapien serine protease catalytic domain
US-09-386-642-8

```

DD 859 GTG 861

RESULT 10  
US-09-518-046-3  
; Sequence 3, Application US/09518046  
; Patent No. 6294663  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Underwood, Lowell J.  
; TITLE OF INVENTION: Transmembrane Serine Protease Overexpressed  
; TITLE OF INVENTION: in Ovarian Carcinoma and Uses Thereof  
; FILE REFERENCE: D6192CIP  
; CURRENT APPLICATION NUMBER: US/09/518,046  
; CURRENT FILING DATE: 2000-03-02  
; EARLIER APPLICATION NUMBER: 09/261,416  
; EARLIER FILING DATE: 1999-03-03

[illegible]

```
Query Match: 26.56% Indels: 38
DB: 4 Gaps: 13
US-09-735-713A-2 (1-306) x US-09-004-731-32 (1-1200)
QY 9 ArgAsnLysLeuLeuLeuLeuGlyIleValPheGluArgGlyLysSerAlaAa 28
DB 366 CGGAATGATCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 410
QY 29 LeuSerLeuProLysAlaProSerCysGlyGlnSerLeuValLysValGlnProTrp-As 48
DB 411 GCGCAGCACCA-----GGCGCTCTGCAGAACGC---AACTTGTGGAGA 452
QY 48 nTyrPheAsnIlePheSerArgIleLeuGlyGlnSerGlnValGluLysGlySerTrp 68
DB 453 ATTGTACACAGGCTCTAATAGATCTAGAGGTCTATTCACAGGATTCGGGTCTCATCC 512
QY 68 oTrpGlnValSerLeuLysGln-----ArgGlnLysHisIleCysGlyGlySe 84
DB 513 TTGGCAGGCGGCTTTCATCAATCAGGATTTTGTAGTAAAAAATTTATCTTGGCGTGGCGC 572
QY 84 rIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsnArgAsnIleVa 104
DB 573 TTTAGTTAGCATCGATGGGTATATACTGCTGCACATTCGTTGCCACG---ACACCAA 629
QY 104 lSerThrLeuAsnValThrAlaGlyGluTrpAspLeuSerGlnThrAspPro-----G1 122
DB 630 TTCGAACCTGAGGTCGATGGGCAATGGGACGTCGCGGACCATGATGAGCGACTGAA 689
QY 122 yGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLysLysPr 142
DB 690 CCACGAGGAATACGCAATCGAACGCAAGAGTTTCATCTTCATATTCATCA-----CC 740
QY 142 oMetAspTrp-----AspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHi 160
DB 741 AACCGATTTCCGGAATGATGATGACCTTAGTGAACCTCGATAGAACTGTTATTTCAAACA 800
QY 160 sPheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCy 180
DB 801 ACATATTTACCTGCTGCTTACTCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 857
QY 180 sThrThrAlaGlyTrpGlyArgLeuThrGluGly---GlyValLeuSerGlnValLeuGl 199
DB 858 AACAGTCGCGGATGGGACGAGGACGAGGACGAGGACGAGGACGAGGACGAGGACGAGGACG 917
QY 199 nGluValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeu-- 218
DB 918 AGAAGTCGATGTCGAGGTGATTCGGAATGAAAGATGCCAGAGGTGTTCCGCTGCTCGGG 977
QY 219 -LysArgProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAs 238
DB 978 TCACAGAGAACCATTCACAGTCTCTCTCTCGCGCGGATATAAAGAGGGTGGTCTGTA 1037
QY 238 pAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpTh 258
DB 1038 TTCTACCAAGGATGATCTGGAGTCTCTAATAATGACAG---ATTGAGGGTAGAAGGAC 1094
QY 258 rLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValAr 278
DB 1095 CCTTGTGGGTCTAGTTCTTGGGGCATTCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1134
QY 278 gLysSerAspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIl 298
DB 1135 -----GAGCATTTACAGGCGGTATATACCAATATACAAAATTCATACCGGTGAT 1184
QY 298 e 298
DB 1185 C 1185
```

## RESULT 12

US-09-004-731-34/c

; Sequence 34, Application US/09004731

; Patent No. 6177258

```
GENERAL INFORMATION:
APPLICANT: Wu Hunter, Shirley
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/004,731
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/749,699
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 1200 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-004-731-34
```

```
Alignment Scores:
Pred. No.: 1,35e-39 Length: 1200
Score: 434.50 Matches: 110
Percent Similarity: 53.16% Conservative: 50
Best Local Similarity: 36.54% Mismatches: 103
Query Match: 26.56% Indels: 38
DB: 4 Gaps: 13
```

```
US-09-735-713A-2 (1-306) x US-09-004-731-34 (1-1200)
QY 9 ArgAsnLysLeuLeuLeuLeuGlyIleValPheGluArgGlyLysSerAlaAa 28
DB 835 CGGAATGATCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 791
QY 29 LeuSerLeuProLysAlaProSerCysGlyGlnSerLeuValLysValGlnProTrp-As 48
DB 790 GCGCAGCACCA-----GGCGCTCTGCAGAACGC---AACTTGTGGAGA 749
QY 48 nTyrPheAsnIlePheSerArgIleLeuGlyGlnSerGlnValGluLysGlySerTrp 68
DB 748 ATTGTACACAGGCTCTAATAGATCTAGAGGTCTATTCACAGGATTCGGGTCTCATCC 689
QY 68 oTrpGlnValSerLeuLysGln-----ArgGlnLysHisIleCysGlyGlySe 84
DB 688 TTGGCAGGCGGCTTTCATCAATCAGGATTTTGTAGTAAAAAATTTATCTTGGCGTGGCGC 629
QY 84 rIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsnArgAsnIleVa 104
DB 628 TTTAGTTAGCATCGATGGGTATATACTGCTGCACATTCGTTGCCACG---ACACCAA 572
QY 104 lSerThrLeuAsnValThrAlaGlyGluTrpAspLeuSerGlnThrAspPro-----G1 122
```

```
Db 571 TTCGAACCTGAAGTGGCGATGGCGAGTCCGCGCAGCAGCGACTGAA 512
Qy 122 yGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLysPr 142
Db 511 CCACGAGGAATACCAATCGAACCAAGAAGATTTCATCTTCATATCA-----CC 461
Qy 142 oMetAspTyr-----AspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyH1 160
Db 460 AACCGATTCCGGATGATGCTAGCTAGTGAACATCGATAGAACTGTATTTCARACA 401
Qy 160 sPheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCy 180
Db 400 ACATATTTTACCTGCTGCTACCTCAT- --AAGCAAAATGAACCTGCGTCAAAAATGCG 344
Qy 180 sThrAlaGlyTrpGlyArgLeuThrGluGly- --GlyValLeuSerGlnValLeuG1 199
Db 343 AACAGTCCCGGATGGGACGAGCAGCGACGCGCAGAGCAGCTGTGCGGCTGCTTACA 284
Qy 199 nGluValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeu- 218
Db 283 AGAAGTCGATCGCAGGTGATTCGGAATGAAGATGCCAGAGGTGTTCCGTGCTGCGGG 224
Qy 219 -LysArgProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAs 238
Db 223 TCGACGAGAAACCAATTCACGATGCTTTCTCTGCGCGCGATATAAAGAGGTGCTCGTGA 164
Qy 238 pAlaCysGlnGlyAspSerGlyLysLeuMetCysArgAsnLysLysGlyAlaTrpTh 258
Db 163 TTCATGCCAAGGTATCTGAGGTCTCTTAATATGCGAG- --ATTGAGGTGTAAGGAC 107
Qy 258 rLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnAsnValar 278
Db 106 CTTGTGGGTCTAGTTCTTGGGCAATGGATGTGGTCT----- 67
Qy 278 gLysSerAspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpI1 298
Db 66 -----GAGCATTTACCAGCGGTATATACCAATATACAAAATTCATACCGTGAT 17
Qy 298 e 298
Db 16 C 16

RESULT 13
US-08-749-699-32
; Sequence 32, Application US/08749699
; Patent No. 6210920
; GENERAL INFORMATION:
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES AND USES THEREOF
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/749,699
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J
; REGISTRATION NUMBER: 32,020
```

```
; REFERENCE/DOCKET NUMBER: 2618-25-C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1200 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1200
US-08-749-699-32

Alignment Scores:
Pred. No.: 1.35e-39 Length: 1200
Score: 434.50 Matches: 110
Percent Similarity: 53.16% Conservative: 50
Best Local Similarity: 36.54% Mismatches: 103
Query Match: 26.56% Indels: 38
DB: 4 Gaps: 13

US-09-735-713A-2 (1-306) x US-08-749-699-32 (1-1200)
Qy 9 ArgAsnLysLeuIleLeuLeuGlyIleValPheGluArgGlyLysSerAlaAla 28
Db 366 CGGAATGATCTGGTCTGCTCGCT-----CGACAGGACATTCGGCCTGA 410
Qy 29 LeuSerLeuProLysAlaProSerCysGlyGlnSerLeuValLysValGlnProTrp-As 48
Db 411 GCCGCAGCACCA-----GGGCGCTCTGCAGAACGC- --AAGTTGTGGAGA 452
Qy 48 nTyrPheAsnIlePheSerArgIleLeuGlyGlySerClnValGluLysGlySerTyrPr 68
Db 453 ATTGTACACGAGGTCTAATAGATCGTAGGAGGTTCATCAACAGGATTCGGGTCTCATCC 512
Qy 68 oTrpGlnValSerLeuLysGln-----ArgGlnLysHisIleCysGlyGlyLys 84
Db 513 TTGGCAGCGCGCTTGATCAAAATCAGGATTTTGTAGTAAAAAATATCTTCGGGTGGCGC 572
Qy 84 rIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsnArgAsnIleVa 104
Db 573 TTTAGTTAGCGATCGATGGGTTATAACTGCTGCACATTGCTGTCACG- --ACACCAAA 629
Qy 104 lSerThrLeuAsnValThrAlaGlyGluTyAspLeuSerGlnThrAspPro-----G1 122
Db 630 TTCGAACCTGAAGTGGCGATTTGGCGCAATGGGACGTCGCGCAGCAGCGACTGAA 689
Qy 122 yGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLysLysPr 142
Db 690 CCACGAGGAATACCAATCGAACCGACGAGCAGCGCAGCAGCTGCTTCATATCA-----CC 740
Qy 142 oMetAspTyr-----AspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyH1 160
Db 741 AACCGATTTCGGGAATGATGCTTAGTCAAACTCGATAGAACTGTTATTTTCAACAA 800
Qy 160 sPheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCy 180
Db 801 ACATATTTTACCTGCTGCTTACCTCAT- --AAGCAAAATGAACCTGCGTCAAAAATGCG 857
Qy 180 sThrAlaGlyTrpGlyArgLeuThrGluGly- --GlyValLeuSerGlnValLeuG1 199
Db 858 AACAGTCCCGGATGGGACGAGCAGCGCAGCAGCAGCTGTCGGCGTCTTACA 917
Qy 199 nGluValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeu- 218
Db 918 AGAAGTCGATCGCAGGTGATTCGAATGAAGATGCCAGAGGTGTTCCGTGCTGCGGG 977
Qy 219 -LysArgProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAs 238
Db 978 TCGACGAGAAACCAATTCACGATGCTTTTCTCTGCGCGGATATAAAGAGGTGCTCGTGA 1037
```



ADDRESSEE: Sheridan Ross P.C.  
STREET: 1700 Lincoln Street, Suite 3500  
CITY: Denver  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80203  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/004,729  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/749,699  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Connell, Gary J.  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2618-25-C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1200 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1200

US-09-004-729-32

Alignment Scores:  
Pred. No.: 1.35e-39 Length: 1200  
Score: 434.50 Matches: 110  
Percent Similarity: 53.16% Conservative: 50  
Best Local Similarity: 36.54% Mismatches: 103  
Query Match: 26.56% Indels: 38  
DB: 4 Gaps: 13

US-09-735-713A-2 (1-306) x US-09-004-729-32 (1-1200)

QY 9 ArgAsnLysLeuIleLeuLeuGlyIleValPhePheGlyLysSerAlaAla 28  
DB 366 CGAATGATCTGGTCTGCTGGT-----CGACAGGAGCATTCGGCCTGA 410  
QY 29 LeuSerLeuProLysAlaProSerCysGlyGlnSerLeuValLysValGlnProTrp-As 48  
DB 411 GCCGACGACCA-----GGCCCTCTGCAGAACGC---AACTTGTGGAGA 452  
QY 48 nTyrPheAsnIlePheSerArgIleLeuGlySerGlnValGluLysGlySerTyPr 68  
DB 453 ATTGTACAGAGGTCTAATAGAATCGTAGAGGTCTCATTCACAGGATTCGGGTCTCATCC 512  
QY 68 cTrpGlnValSerLeuLysGln-----ArgGlnLysHisIleCysGlyClySe 84  
DB 513 TTGGCAGCGCGCTTTGATCAATCATCAGATTTTGTAGTAAAAAATTATCTTCGGTGGCGC 572  
QY 84 rIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsnArgAsnIleVa 104  
DB 573 TTTAGTAGCGATCGATGGTTATACGTCTGCACATTCGGTGGCAGC---ACACAAA 629  
QY 104 lSerThrLeuAsnValThrAlaGlyClyTrpAspLeuSerGlnThrAspPro-----G1 122  
DB 630 TTCGAACCTGAAGGTGCGATTGGCGAATGGGACGTCCGCGACACGATGACGACTGAA 689  
QY 122 yGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLysLysPr 142  
DB 122 yGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLysLysPr 142

DB 690 CCACGAGGAATACGCAATCGAAGCGAAAGAGTTTCATCTTCATATTC-----CC 740  
QY 142 oMetAspTyr-----AspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHi 160  
DB 741 AACCGATTTCCGGAATGATGAGCTTAGTGAACCTCGATAGAACTGTTATTTTCAACA 800  
QY 160 sPheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCy 180  
DB 801 ACATATTTTACCTGCTGCTTACCTCAT---AAGCAATGAAACTGGCTGGAATAATGCG 857  
QY 180 sThrThrAlaGlyTrpGlyArgLeuThrGluGly---GlyValLeuSerGlnValLeuGl 199  
DB 858 AACAGTCGCGGATGGGACGAGCAGGACGAGCAGCTGTGCGCGCTGTCTTACA 917  
QY 199 nGluValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeu-- 218  
DB 918 AGAAGTCGATGTCGAGGTGATTCGGAATGAAAGATGCCAGAGGTGTCGCTGCGGG 977  
QY 219 -LysArgProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAs 238  
DB 978 TCGACGAGAAACCAATTCACGATGCTCTCTGCGCGCGATATAAAGAGGTGTCGTGA 1037  
QY 238 pAlaCysGlnGlyAspSerGlySerLeuMetCysArgAsnLysLysGlyAlaTrpTh 258  
DB 1038 TTCATGCCAAGGTGATTCGAGGTCCTCTAATAATGCAG---ATTGAGGGTAGAAGGAC 1094  
QY 258 rLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValAr 278  
DB 1095 CTTTGGGTCTAGTTCTTGGGGCATTCGATGTCGT-----1134  
QY 278 gLysSerAspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrp1 298  
DB 1135 -----GAGCATTTACCAGGCGTATATACCAATATACAAAAATTCATACCGTGGAT 1184  
QY 298 e 298  
DB 1185 C 1185

Search completed: March 25, 2003, 02:04:14  
Job time : 66 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 03:48:49 ; Search time 20 Seconds  
(without alignments)

634.588 Million cell updates/sec

Title: US-09-735-713A-2

Perfect score: 1636

Sequence: 1 MSUKMLISRNKLLLLGIVF.....IFTDISKVLSWIHEHIQTGN 306

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	486	29.7	786	1	Q05319 drosophila
2	466	28.5	638	1	P03952 homo sapien
3	464	28.4	638	1	P26262 mus musculus
4	456.5	27.9	454	1	P57727 homo sapien
5	447.5	27.4	638	1	P14272 rattus norv
6	446	27.3	342	1	Q9esd1 mus musculus
7	446	27.3	342	1	Q9es87 rattus norv
8	444	27.1	343	1	P16651 homo sapien
9	442.5	27.0	417	1	Q05981 homo sapien
10	442.5	27.0	1035	1	P98072 bos taurus
11	440.5	26.9	416	1	Q35453 mus musculus
12	438.5	26.8	416	1	Q05511 rattus norv
13	438.5	26.8	1034	1	P98074 sus scrofa
14	438	26.8	457	1	Q9h383 homo sapien
15	433	26.5	593	1	P98140 bos taurus
16	431.5	26.4	1019	1	P98073 homo sapien
17	429	26.2	290	1	Q9bqr3 homo sapien
18	427	26.1	311	1	Q9qul7 mus musculus
19	426	26.0	855	1	P56677 mus musculus
20	425	26.0	1069	1	P97435 mus musculus
21	423	25.9	455	1	Q9er04 mus musculus
22	421.5	25.8	317	1	Q9gzn4 homo sapien
23	421.5	25.8	321	1	Q9nrr2 homo sapien
24	419.5	25.6	237	1	P00765 astacus flu
25	419	25.6	264	1	P40313 homo sapien
26	419	25.6	273	1	Q9xsm2 ovis aries
27	418	25.6	306	1	Q9er10 mus musculus
28	418	25.6	625	1	P03951 homo sapien
29	415.5	25.4	275	1	P35037 anopheles g
30	414.5	25.3	615	1	P00748 homo sapien
31	413.5	25.3	263	1	P04813 canis fami
32	413.5	25.3	274	1	TRY1_ANOGA
33	413	25.2	437	1	Q9nrs4 homo sapien

#### ALIGNMENTS

##### RESULT 1

ID	STUB_DROME	STANDARD;	PRT;	786 AA.
AC	Q05319;			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Serine proteinase stubble (EC 3.4.21.-) (Stubble-stubbloid protein).			
GN	SB OR SB-SBD.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;			
OC	Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;			
OC	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Oregon-R;			
RX	MEDLINE=93281671; PubMed=7685111;			
RA	Appel L.F., Prout M., Abu-Shumays R., Hammonds A., Garbe J.C.,			
RA	Fristrom D., Fristrom J.;			
RT	"The Drosophila Stubble-stubbloid gene encodes an apparent			
RT	transmembrane serine protease required for epithelial			
RT	morphogenesis.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 90:4937-4941(1993).			
CC	-!- FUNCTION: HORMONE DEPENDENT PROTEASE REQUIRED FOR EPITHELIAL			
CC	MORPHOGENESIS. HAS A DUAL FUNCTION, DETACHES IMAGINAL DISC CELLS			
CC	FORM EXTRACELLULAR MATRICES THROUGH ITS EXTRACELLULAR PROTEOLYTIC			
CC	DOMAIN AND TRANSMITS AN OUTSIDE-TO-INSIDE SIGNAL TO ITS			
CC	INTRACELLULAR DOMAIN TO MODIFY THE CYTOSKELETON DURING			
CC	MORPHOGENESIS. MAY BE ABLE TO ACTIVATE ITSELF.			
CC	-!- SUBCELLULAR LOCATION: Type II membrane protein.			
CC	-!- INDUCTION: BY 20-HYDROXYECYDSONE (20HE).			
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			
CC	-!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-24 IS THE INITIATOR.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; L11451; AAA28918.1; -			
DR	HSSP; P00763; 1DPO.			
DR	MEROPS; S01.225; -			
DR	FlyBase; FBgn0003319; Sb.			
DR	InterPro; IPR001314; Chymotrypsin.			
DR	InterPro; IPR001254; Ser_protease_Try.			
DR	Pfam; PF00089; trypsin.1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	SMART; SM00020; Tryp_Spc; 1.			
DR	PROSITE; PS00240; TRYPSIN_DOM; 1.			
DR	PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.			
DR	PROSITE; PS00135; TRYPSIN_SER; 1.			
KW	Hydrolase; Serine protease; Transmembrane; Glycoprotein; Zymogen;			

34	409	25.0	263	1	CTRB_HUMAN	PI7538 homo sapien
35	409	25.0	275	1	TRY4_ANOGA	P35038 anopheles g
36	408.5	25.0	415	1	ACRO_PIG	P08001 sus scrofa
37	408.5	25.0	603	1	FAI2_CAVPO	Q04962 cavia porce
38	408	24.9	490	1	TMS2_MOUSE	Q9jiq8 mus musculu
39	407.5	24.9	459	1	PRTC_PIG	Q9qlp2 sus scrofa
40	407	24.9	324	1	TEST_MOUSE	Q9jhi7 mus musculu
41	407	24.9	492	1	TMS2_HUMAN	O15393 homo sapien
42	406	24.8	254	1	TRY3_AEDAE	P29786 aedes aegyp
43	404	24.7	407	1	FA7_BOVIN	P22457 bos taurus
44	403	24.6	314	1	TEST_HUMAN	Q9y6m0 homo sapien
45	402.5	24.6	275	1	TRYT_PIG	Q9n2dl sus scrofa



```
FT CHAIN 391 638 PLASMA KALLIKREIN LIGHT CHAIN.
FT DOMAIN 20 105 APPLE 1.
FT 110 195 APPLE 2.
FT DOMAIN 200 285 APPLE 3.
FT DOMAIN 291 376 APPLE 4.
FT DOMAIN 389 621 SERINE PROTEASE.
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .).
FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .).
FT CARBOHYD 494 494 N-LINKED (GLCNAC. . .).
FT ACT_SITE 434 434 ACT_SITE
FT ACT_SITE 483 483 ACT_SITE
FT ACT_SITE 578 578 ACT_SITE
FT DISULFID 21 104 CHARGE RELAY SYSTEM.
FT DISULFID 47 77 CHARGE RELAY SYSTEM.
FT DISULFID 51 57
FT DISULFID 111 194
FT DISULFID 137 166
FT DISULFID 141 147
FT DISULFID 201 284
FT DISULFID 227 256
FT DISULFID 231 237
FT DISULFID 292 375
FT DISULFID 318 347
FT DISULFID 322 328
FT DISULFID 340 345
FT DISULFID 383 503
FT DISULFID 419 435
FT DISULFID 517 584
FT DISULFID 548 563
FT DISULFID 574 602
FT VARIANT 143 143
FT VARIANT 202 202 N -> S (COMMON POLYMORPHISM).
FT VARIANT 208 208 /FTid-Var_013598.
FT VARIANT 208 208 H -> Q.
FT VARIANT 208 208 /FTid-Var_013599.
FT VARIANT 208 208 H -> P.
FT VARIANT 208 208 /FTid-Var_013600.
FT SEQUENCE 638 AA; 71369 MW; E62F9C1053838FB4 CRC64;

Query Match 28.5%; Score 466; DB 1; Length 638;
Best Local Similarity 35.5%; Pred. No. 9.4e-36;
Matches 91; Conservative 51; Mismatches 96; Indels 18; Gaps 5;

QY 54 SRILGGQVEKSGVPMQVSLKQR---QKHLGGSGIVSPQVITAACHIANRNIVSTLNVT 110
DB 389 TRIVGGTNSNGEMPMQVSLQVLTAAQHLCGGSLGHQWLVTAACHFDGLPLQDVWRIY 448
QY 111 AGEYDLSQTDPRGEQTLTIETVTHPHFSTKKPMYDIALKMGAGAFQGHVGPICLPEL 170
DB 449 SGILNLSDIITKDPFQSIKEIIHQNYKVSEG-NHDIALIKLQAPLNYTEFQRPICLPSK 507
QY 171 REQFEAGFICTTAGBRLTEGGVLSQVLEVNLPILTWEECVAAALLFLKRPISGKTEFLCT 230
DB 508 GDTSTIYTNWVTGWGFSKEKEIQNLQKVNPLVTNEECQKRYQDYK---ITQRMVCA 564
QY 231 GPPDGGDCAQDGGSLMCRNKKAGWTLAGVTSWGLCGCGRWNNVKKDQSGSPGLFTD 290
DB 565 GYEGKGKDGKGGGGLPVCKH-NGMWRVLGITSWEGGCAR-----REQPGVYTK 613
QY 291 ISKVLWISHEIQTGN 306
DB 614 VAEMDWILEKTQSSD 629

RESULT 3
ID KAL_MOUSE STANDARD; PRT; 638 AA.
AC P26262;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasma kallikrein precursor (EC 3.4.21.34) (Plasma prekallikrein)
```

```
DE (Kininogenin) (Fletcher factor).
GN KLB1 OR KLB3 OR PK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=BAIB/G; TISSUE=Liver;
RX MEDLINE=91090844; PubMed=2264928;
RA Seidah N.G., Sawyer N., Hamelin J., Mion P., Beaubien G.,
RA Brachpapa L., Rochemont J., Mbikay M., Chretien M.;
RT "Mouse plasma kallikrein: cDNA structure, enzyme characterization,
RT and comparison of protein and mRNA levels among species.";
RL DNA Cell Biol. 9:737-748(1990)
CC -|- FUNCTION: THE ENZYME CLEAVES LYS-ARG AND ARG-SER BONDS. IT
CC ACTIVATES, IN A RECIPROCAL REACTION, FACTOR XII AFTER ITS BINDING
CC TO A NEGATIVELY CHARGED SURFACE. IT ALSO RELEASES BRADYKININ FROM
CC HMW KININOGEN AND MAY ALSO PLAY A ROLE IN THE RENIN-ANGIOTENSIN
CC SYSTEM BY CONVERTING PRORENIN INTO RENIN.
CC -|- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Xaa and Lys-|-Xaa
CC bonds, including Lys-|-Arg and Arg-|-Ser bonds in (human)
CC kininogen to release bradykinin.
CC -|- SUBUNIT: THE ZYMOGEN IS ACTIVATED BY FACTOR XIIA, WHICH CLEAVES
CC THE MOLECULE INTO A LIGHT CHAIN, WHICH CONTAINS THE ACTIVE SITE,
CC AND A HEAVY CHAIN, WHICH ASSOCIATES WITH HMW KININOGEN. THESE
CC CHAINS ARE LINKED BY ONE OR MORE DISULFIDE BONDS.
CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMA KALLIKREIN
CC SUBFAMILY.
CC -|- SIMILARITY: CONTAINS 4 APPLE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M58588; AAA63393.1; -.
DR PIR; A36557; KOMSPL.
DR HSP; P00750; LRTF.
DR MEROPS; S01.212; -.
DR MGD; MGI:102849; KLB3.
DR InterPro; IPR000177; Apple.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00024; PAN; 4.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00005; APPLEDOMAIN.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00223; APPLE; 4.
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS00495; APPLE; 4.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolyase; Serine protease; Glycoprotein; Plasma; Zymogen; Signal;
KW Fibrinolysis; Blood coagulation; Inflammatory response; Liver;
KW Repeat..
KW SIGNAL 1 19
FT CHAIN 20 390 PLASMA KALLIKREIN HEAVY CHAIN.
FT CHAIN 391 638 PLASMA KALLIKREIN LIGHT CHAIN.
FT DOMAIN 20 105 APPLE 1.
FT DOMAIN 110 195 APPLE 2.
FT DOMAIN 200 285 APPLE 3.
FT DOMAIN 291 376 APPLE 4.
FT DOMAIN 389 621 SERINE PROTEASE.
FT DOMAIN 127 127 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (PROBABLE).
FT CARBOHYD 453 453 O-LINKED (PROBABLE).
```

```

FT CARBOHYD 494 N-LINKED (GLCNAC... ) (PROBABLE).
FT ACT_SITE 434 CHARGE RELAY SYSTEM.
FT ACT_SITE 483 CHARGE RELAY SYSTEM.
FT ACT_SITE 578 CHARGE RELAY SYSTEM.
FT DISULFID 21 BY SIMILARITY.
FT DISULFID 47 BY SIMILARITY.
FT DISULFID 51 BY SIMILARITY.
FT DISULFID 111 BY SIMILARITY.
FT DISULFID 137 BY SIMILARITY.
FT DISULFID 141 BY SIMILARITY.
FT DISULFID 201 BY SIMILARITY.
FT DISULFID 227 BY SIMILARITY.
FT DISULFID 231 BY SIMILARITY.
FT DISULFID 292 BY SIMILARITY.
FT DISULFID 318 BY SIMILARITY.
FT DISULFID 322 BY SIMILARITY.
FT DISULFID 340 BY SIMILARITY.
FT DISULFID 383 BY SIMILARITY.
FT DISULFID 419 BY SIMILARITY.
FT DISULFID 517 BY SIMILARITY.
FT DISULFID 548 BY SIMILARITY.
FT DISULFID 574 BY SIMILARITY.
SQ SEQUENCE 638 AA; 71368 MW; CC27C93AB1086599 CRC64;

Query Match 28.4%; Score 464; DB 1; Length 638;
Best Local Similarity 36.8%; Pred. No. 1.4e-35;
Matches 95; Conservative 43; Mismatches 102; Indels 18; Gaps 6;

QY 52 IFSRIILGSOVEKGYPMQVSLKOR---QKHICGGSIYSPOWITAAHCIAANRNIVSTLN 108
DB INARIVGGTNASLGEPWQVSLQVLSQTHLGGSIIGROWLFAAHCDFGIPDPVWR 446

QY 109 VTAGEYDLSQTDGPGQTITETVTHPHFSTKKPKMDYDIALKMGAFQFGHFVGPICLP 168
DB IGLSILSEITKPKPSSRIKELIITHQYKVSSEG-NYDIALIKLQTLPLNYTEFQKPICLP 505

QY 169 ELREQFEAGFICTTACGRLTEGGVLSQVLEVNPLPILTBECVAAALTLKRPISGKFTL 228
DB SKADNTNTYTNCWVTGWYTKQGGTQNLQKATLPNPECQK---KYRDYVINKQMI 562

QY 229 CTGFPGDGRDAGQSGSLMCRNKKGAWTLAGVTSWGLGCGRGWRNVRKSDQSGPIF 288
DB CAGYKEGTDACKDGGGLVCKH-SGRWLVGITSWEGGCR-----KDQ--PGVY 611

QY 289 TDISKVLNWIHGIOTGN 306
DB TKVSEYMDWILEKTSQSD 629

RESULT 4
TMS3_HUMAN
ID TMS3_HUMAN STANDARD; PRT; 454 AA.
AC P57727;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transmembrane protease, serine 3 (EC 3.4.21.-) (Serine protease
TADG-12) (Tumor associated differentially-expressed gene-12 protein).
GN TMRPSS3 OR TADG12 OR ECHOS1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A. (ISOFORMS A AND TRUNCATED).
RC TISSUE=Ovarian carcinoma;
RX MEDLINE=20521358; PubMed=11068177;
RA Underwood L.J., Shigemasa K., Tanimoto H., Beard J.B., Schneider E.N.,
RA Wang Y., Parmley T.H., O'Brien T.J.;
RT "Ovarian tumor cells express a novel multi-domain cell surface serine
RT protease.";
RL Biochim. Biophys. Acta 1502:337-350(2000).
RN [2]

```

```

RP SEQUENCE FROM N.A. (ISOFORMS A; B/C AND D), AND VARIANT ILE-53.
RX MEDLINE=20578749; PubMed=11137999;
RA Scott H.S., Kudoh J., Wattenhofer M., Shibuya K., Berry A., Chraast R.,
RA Guipponi M., Wang J., Kawasaki K., Asakawa S., Minoshima S.,
RA Younus F., Mehdi S.Q., Radhakrishna U., Papasavvas M.P., Gehrig C.,
RA Rossier C., Korostishevsky M., Gal A., Shimizu N., Bonne-Tamir B.,
RA Antonarakis S.E.;
RT "Insertion of beta-satellite repeats identifies a transmembrane
RT protease causing both congenital and childhood onset autosomal
RT recessive deafness.";
RL Nat. Genet. 27:59-63(2001).
RN [3]
RP VARIANTS DFBN8/DFNB10 CYS-251 AND LEU-404.
RX MEDLINE=21354482; PubMed=11462234;
RA Masmoudi S., Antonarakis S.E., Schwede T., Ghorbel A.M., Gratri M.,
RA Pappasavas M.P., Drira M., Elgaied-Boullia A., Wattenhofer M.,
RA Rossier C., Scott H.S., Ayadi H., Guipponi M.;
RT "Novel missense mutations of TMRPSS3 in two consanguineous Tunisian
RT families with non-syndromic autosomal recessive deafness.";
RL Hum. Mutat. 18:101-108(2001).
RN [4]
RP VARIANTS DFBN8/DFNB10 TRP-109; PHE-194 AND ARG-407, AND VARIANTS
RP ILE-53; SER-111 AND VAL-253.
RX PubMed=11424922;
RA Ben-Yosef T., Wattenhofer M., Riazuddin S., Ahmed Z.M., Scott H.S.,
RA Kudoh J., Shibuya K., Antonarakis S.E., Bonne-Tamir B.,
RA Radhakrishna U., Naz S., Ahmed Z., Riazuddin S., Pandya A.,
RA Nance W.E., Wilcox E.R., Friedman T.B., Morell R.J.;
RT "Novel mutations of TMRPSS3 in four DFBN8/B10 families segregating
RT congenital autosomal recessive deafness.";
RL J. Med. Genet. 38:396-400(2001).
RN [5]
RP VARIANT DFBN8/DFNB10 GLY-103, AND VARIANTS ASN-173 AND THR-426.
RX PubMed=11907649;
RA Wattenhofer M., Di Torio V., Rabionet R., Dougherty L., Pampanos A.,
RA Schwede T., Montserrat-Sentis B., Arbones L., Illades T.,
RA Pasquodibiscaglia A., D'Amelio M., Alwan S., Rossier C., Dahl H.-H.M.,
RA Petersen M.B., Estivill X., Gasparini P., Scott H.S.,
RA Antonarakis S.E.;
RT "Mutations in the TMRPSS3 gene are a rare cause of childhood
RT nonsyndromic deafness in Caucasian patients.";
RL J. Mol. Med. 80:124-131(2002).
CC -!- FUNCTION: Probable protease.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC -!- ALTERNATIVE PRODUCTS: 4 ISOFORMS; A (SHOWN HERE), B/C, D AND
CC TRUNCATED/TADG-12V; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: Expressed in many tissues including fetal
CC cochlea. The truncated isoform is found at increased levels in
CC some carcinomas.
CC -!- DISEASE: Defects in TMRPSS3 are a cause of two forms of autosomal
CC neurosensory childhood-onset forms of deafness, DFBN8 and DFBN10.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -!- SIMILARITY: CONTAINS 1 SRCR DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLASS A DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF201380; AAC37012.1; -
DR EMBL; AB038157; BAB20077.1; -
DR EMBL; AB038158; BAB20078.1; -
DR EMBL; AB038159; BAB20079.1; -
DR EMBL; AB038160; BAB20080.1; -
DR HSSP; P00763; 1DPO.
DR MEROPS; S01.079; -.
DR Genew; HGNC:11877; TMRPSS3.
DR MIM; 605511; -.
DR MIM; 601072; -.

```



FT	DISULFID	201	284	BY SIMILARITY.
FT	DISULFID	227	256	BY SIMILARITY.
FT	DISULFID	231	237	BY SIMILARITY.
FT	DISULFID	292	375	BY SIMILARITY.
FT	DISULFID	318	347	BY SIMILARITY.
FT	DISULFID	322	328	BY SIMILARITY.
FT	DISULFID	340	345	BY SIMILARITY.
FT	DISULFID	383	503	BY SIMILARITY.
FT	DISULFID	419	435	BY SIMILARITY.
FT	DISULFID	517	584	BY SIMILARITY.
FT	DISULFID	548	563	BY SIMILARITY.
FT	DISULFID	574	602	BY SIMILARITY.
SQ	SEQUENCE	638 AA;	71273 MH; 454BEB27E8CA9F88 CRC64;	
Query Match				
Best Local Similarity 27.4%; Score 447.5; DB 1; Length 638;				
Matches 97; Conservative 49; Mismatches 118; Indels 27; Gaps				
Qy	18	IVFFERKSA-ALSLKAPSCGSLVKVQPNVFNIFSRILGSGVKEKSGYPQVSLQKR	767	
Db	360	ITYEAQSGSYSLRCLKVWESSDCTTKIN-----ARIVGTSNLSLGEPWQVSLQVK	411	
Qy	77	--OKHICGGSIVSPQWVITAHAHCIANRNIVSTLVNTAGEYDLSDQDPGEQTLTETVII	133	
Db	412	LVSONHHCSSIIIGRWILTAHCFDGPYPDQWRIYGGILNLSEITNTKTFSSIKELII	471	
Qy	134	HPHFSTKKPMYDIALKMGAGAFQGHFVGPICLPRLRQFEAGFICTTAGWGRLTGEGV	193	
Db	472	HOKYKMEG- SYDIALIKLQTLPLNYTEFQKPICLPSKADNTIYNCWVTGHWGTYKERGE	530	
Qy	194	LSQVLQVNPILPTWBECAVALLTLKRPISGKTFCTGTPDGGRCACGDSGSLMCRNK	253	
Db	531	TQNILOKATILPLVPNEBCQ---KYRDYVITQMICAGYKEGGIDACKGDSGGLVCKH	586	
Qy	254	KGAWTLAGVTSWGLGCGRGWRNNRVKSDGSGCIFTDISKVLSWTHEHTQT	304	
Db	587	SGRWQLVITISWEGGCAR-----REQPGVYTKVAEYIDWILEKIQS	627	
RESULT 6				
PS8_MOUSE	ID	PS8_MOUSE	STANDARD;	PRT; 342 AA.
AC	QPSDL;			
DC	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Prostasin precursor (EC 3.4.21.-) (Channel activating protease 1).			
GN	PS88 OR CAP1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
ON	NCBI_TaxID=10090;			
OX	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=20335202; PubMed=10770960;			
RX	Vuagniaux G., Vallet V., Jaeger N.F., Pfister C., Bens M., Farman N.,			
RA	Courtis-Couty N., Vandewalle A., Rossler B.C., Hummler E.;			
RT	"Activation of the aniloride-sensitive epithelial sodium channel by			
RT	the serine protease mCAP1 expressed in a mouse cortical collecting			
RT	duct cell line."			
RL	J. An. Soc. Nephrol. 11:828-834(2000).			
CC	-!- FUNCTION: POSSESSES A TRYPSIN-LIKE CLEAVAGE SPECIFICITY (BY			
CC	SIMILARITY). ACTIVATES AMILORIDE-SENSITIVE SODIUM CHANNELS.			
CC	-!- SUBUNIT: HETERODIMER OF TWO CHAINS, LIGHT AND HEAVY, HELD BY A			
CC	DISULFIDE BOND (BY SIMILARITY).			
CC	-!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. SECRETED AFTER CLEAVAGE OF			
CC	ITS C-TERMINUS (BY SIMILARITY).			
CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.			
CC	-!- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A			
CC	FRAMESHIFT IN POSITION 339.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collabora			
CC	between the SwissInstitute of Bioinformatics and the EMBL outstatio			







CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -!- SIMILARITY: CONTAINS 2 LDL-RECEPTOR CLASS A DOMAINS.  
 CC -!- SIMILARITY: CONTAINS 2 CUB DOMAINS.  
 CC -!- SIMILARITY: CONTAINS 1 SEA DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 SRCR DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 MAM DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U09859; AAB40026.1; -;  
 CC EMBL; L19663; AAB16035.1; -;  
 CC PIR; A61436; A61436.  
 CC HSP; P00763; IDPO.  
 CC MEROPS; S01.156; -;  
 CC InterPro; IPR000859; CUB\_domain.  
 CC InterPro; IPR001314; Chymotrypsin.  
 CC InterPro; IPR002172; LDL\_recept\_A.  
 CC InterPro; IPR000998; MAM\_domain.  
 CC InterPro; IPR000082; SEA\_domain.  
 CC InterPro; IPR001254; Ser\_protease\_Try.  
 CC InterPro; IPR001190; Srcr\_receptor.  
 CC Pfam; PF00057; ldl\_recept\_a; 2.  
 CC Pfam; PF00089; trypsin; 1.  
 CC Pfam; PF00431; CUB; 2.  
 CC Pfam; PF00530; SRCR; 1.  
 CC Pfam; PF00629; MAM; 1.  
 CC Pfam; PF01390; SEA; 1.  
 CC PRINTS; PR00722; CHYMOTRYPSIN.  
 CC SMART; SM00042; CUB; 2.  
 CC SMART; SM00192; LDLA; 2.  
 CC SMART; SM00137; MAM; 1.  
 CC SMART; SM00200; SEA; 1.  
 CC SMART; SM00202; SR; 1.  
 CC SMART; SM00020; Tryp\_SPC; 1.  
 CC PROSITE; PS01180; CUB; 2.  
 CC PROSITE; PS01209; LDLA\_1; 2.  
 CC PROSITE; PS00068; LDLA\_2; 2.  
 CC PROSITE; PS00740; MAM\_1; 1.  
 CC PROSITE; PS00060; MAM\_2; 1.  
 CC PROSITE; PS00024; SEA; 1.  
 CC PROSITE; PS00420; SRCR\_1; FALSE\_NEG.  
 CC PROSITE; PS02087; SRCR\_2; 1.  
 CC PROSITE; PS0240; TRYPSIN\_DOM; 1.  
 CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
 CC Signal-anchor; Glycoprotein; Myristate; Hydrolase;  
 CC Serine protease; Zymogen; Transmembrane; Repeat; Alternative splicing.  
 CC Signal-anchored; CATALYTIC CHAIN (HEAVY CHAIN).  
 CC CATALYTIC CHAIN (LIGHT CHAIN).  
 CC CYTOPLASMIC (POTENTIAL).  
 CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 CC (POTENTIAL).  
 CC EXTRACELLULAR (POTENTIAL).  
 CC DOMAIN 48 1035  
 CC DOMAIN 54 169  
 CC DOMAIN 197 238  
 CC DOMAIN 240 350  
 CC CUB 1.  
 CC MAM.  
 CC CUB 2.  
 CC DOMAIN 540 650  
 CC DOMAIN 657 695  
 CC LDL-RECEPTOR CLASS A 2.  
 CC SRCR.  
 CC DOMAIN 694 787  
 CC DOMAIN 801 1035  
 CC SERINE PROTEASE.  
 CC ACT\_SITE 841 841  
 CC CHARGE RELAY SYSTEM (BY SIMILARITY).  
 CC ACT\_SITE 892 892  
 CC CHARGE RELAY SYSTEM (BY SIMILARITY).  
 CC ACT\_SITE 987 987  
 CC LIPID 2 2  
 CC MYRISTATE (POTENTIAL).  
 CC BY SIMILARITY.  
 CC DISULFID 199 212  
 CC DISULFID 206 225

FT DISULFID 219 236 BY SIMILARITY.  
 FT DISULFID 659 671 BY SIMILARITY.  
 FT DISULFID 666 684 BY SIMILARITY.  
 FT DISULFID 678 693 BY SIMILARITY.  
 FT DISULFID 788 812 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 826 842 BY SIMILARITY.  
 FT DISULFID 926 993 BY SIMILARITY.  
 FT DISULFID 957 972 BY SIMILARITY.  
 FT DISULFID 983 1011 BY SIMILARITY.  
 FT CARBOHYD 116 116 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 170 170 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 194 194 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 233 233 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 264 264 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 404 404 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 456 456 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 486 486 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 519 519 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 550 550 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 646 646 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 698 698 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 722 722 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 741 741 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 762 762 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 864 864 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 903 903 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 965 965 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT VARSPIC 166 192 MISSING (IN SHORT ISOFORM).  
 FT CONFLICT 808 808 R -> Y (IN REF. 3).  
 SQ SEQUENCE 1035 AA; 114887 MW; E207970B08296E13 CRC64;  
 Query Match 27.0%; Score 442.5; DB 1; Length 1035;  
 Best Local Similarity 34.9%; Pred. No. 2.7e-33;  
 Matches 96; Conservative 54; Mismatches 90; Indels 35; Gaps 9;  
 QY 36 SCGQSLV--KVOPWNYFNIFSRILGSGQVEKSYPMOVSVKQRQKHGICGSGIVSPQWVIT 93  
 DB 787 SCGKLVTVQEVSP-----KIVGSDSREGANFWVALYFDQDVCGASVSRDLVS 838  
 QY 94 AAHCIAANRNI-VSTLNVTAGEDYLSQ-TDPGEQTITETVTHPHFSTKPKMDYDIALK 151  
 DB 839 AAHCYVYGRNMEPSKWKAVLGLHMASNLSPQIETRLIDQIVINPHYN-KRRKNNDIATMMH 897  
 QY 152 MAGAFQFGHEVGPICLPRLREOFEAGFICTTAGWGRLEGGVLSOVLOEVNLPILTWEBC 211  
 DB 898 LEMKYNVTDYIQPICLPEENQVFPFGRICISAGWALYIQGSTADVLDQADVPLLSNEKC 957  
 QY 212 VAAL----LTLKRPISGKTFCTGPDGGRDACQDGGSLMCRNKKGAWTLGAVTSMGL 267  
 DB 958 QQQMPEYNIT-----ENMVCAGYEAGVDSCQDGGPLMCQ-ENNRWLLAGVTSGFY 1009  
 QY 268 GCGRWNRNVRKSDGSGFGIFTDISKVLWSHEHI 302  
 DB 1010 QCALPNR-----ECVYARVPRFTEWIOSFL 1034  
 RESULT 11  
 HEPS\_MOUSE  
 ID HEPS\_MOUSE STANDARD; PRT; 416 AA.  
 AC O35453;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Serine protease hepsin (EC 3.4.21.-).  
 GN HPN.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.



```
Query Match      26.8%; Score 438.5; DB 1; Length 416;
Best Local Similarity 33.0%; Pred. No. 2.le-33;
Matches 102; Conservative 48; Mismatches 112; Indels 47; Gaps 11;

QY 11 KLILLGLVFFERGSALSLPKAPSCGSLVKVQPNWYFISRLILGSSQVEXGSPWQ 70
Db 129 RLDDVISVCDPRFLTAT---CODGRRKLPV-----DRVGGQSSLGWRPWQ 176
QY 71 VSLKQRKHGGGSIQVQWVITAAHCIAANRN-IVSTLNVTAGEDYDLSDTDPGEOTLIE 129
Db 177 VSLRDGTHLCGSLSGDWLTAACFPERNRLVSRWRVAGA--VARTSPHAFVLGVQ 234
QY 130 TVIIHPH-----STKKPMVDYDIALKMGAFQFGHFVGPICLPDLRPFQFAGFICTTAG 184
Db 235 AVIYHGGVLPFRDPTIDENSNDIALVHLSSPLPLEYIQPVCLPAAGQALVDGKVCVTG 294
QY 185 WRLTEGGVLSQVLOEVNLPILTWECEVAALLTLKRP-----ISGKFLCTGFPDGGRD 238
Db 295 WNTGYQQAQVQLQEARVPIISNEVC-----NSPDFYGNQIKPKMF-CAGYPEGGID 346
QY 239 ACOGDSGSGSLMCRNK---KGAWTLAGVTSWGLGCGRGNRNVRKSDQSGPFIETDISKVL 295
Db 347 ACOGDSGGHFVEDRISGTSRMLCGIIVSWGTCAL-----ARK-----PGVYTKVIDFR 396
QY 296 SWIHEHIQT 304
Db 397 EWIFQAIKT 405

RESULT 13
ENTK_PIG
ID ENTK_PIG STANDARD; PRT; 1034 AA.
AC P98074;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enteropeptidase precursor (EC 3.4.21.9) (Enterokinase).
GN PRS7 OR ENTK.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Duodenal mucosa;
RA MEDLINE=94327548; PubMed=8051081;
RA Matsushima M., Ichinose M., Yahagi N., Kakei N., Tsukada S.,
RA Miki K., Kurokawa K., Tashiro K., Shiokawa K., Shinomiya K.,
RA Uneyama H., Inoue H., Takahashi T., Takahashi K.;
RA "Structural characterization of porcine enteropeptidase.";
RL J. Biol. Chem. 269:19976-19982(1994).
CC A) IF FUNCTION: RESPONSIBLE FOR INITIATING ACTIVATION OF PANCREATIC
CC PROTEOLYTIC PROENZYMES (TRYPSIN, CHYMOTRYPSIN AND CARBOXYPEPTIDASE
CC A). IT CATALYZES THE CONVERSION OF TRYPSINOGEN TO TRYPSIN WHICH IN
CC TURN ACTIVATES OTHER PROENZYMES INCLUDING CHYMOTRYPSINOGEN,
CC PROCARBOXYPEPTIDASES, AND PROELASTASES.
CC -I- CATALYTIC ACTIVITY: Selective cleavage of 6-Lys-Ile-7 bond in
CC trypsinogen.
CC -I- SUBUNIT: HETEROTRIMER OF A CATALYTIC (LIGHT) CHAIN, A MULTIDOMAIN
CC (HEAVY) CHAIN, AND A MINI CHAIN.
CC -I- SUBCELLULAR LOCATION: Type II membrane protein (Probable).
CC -I- PTM: THE CHAINS ARE DERIVED FROM A SINGLE PRECURSOR THAT IS
CC CLEAVED BY A TRYPSIN-LIKE PROTEASE.
CC -I- PTM: THE MINI CHAIN MAY BE CLEAVED BY ELASTASE.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -I- SIMILARITY: CONTAINS 2 LDL-RECEPTOR CLASS A DOMAINS.
CC -I- SIMILARITY: CONTAINS 2 CUB DOMAINS.
CC -I- SIMILARITY: CONTAINS 1 SEA DOMAIN.
CC -I- SIMILARITY: CONTAINS 1 SRCR DOMAIN.
CC -I- SIMILARITY: CONTAINS 1 MAM DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D30799; BAA06459.1; -
DR HSSP; P00763; 1DPO.
DR MEROPS; S01.156; -.
DR InterPro; IPR000859; CUB_domain.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR002172; LDL_recept_A.
DR InterPro; IPR000998; MAM_domain.
DR InterPro; IPR000082; SEA_domain.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00057; ldl_recept_a; 2.
DR Pfam; PF00089; trypsin; 1.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00530; SRCR; 1.
DR Pfam; PF00629; MAM; 1.
DR Pfam; PF01390; SEA; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00042; CUB; 2.
DR SMART; SM00192; LDLa; 2.
DR SMART; SM00137; MAM; 1.
DR SMART; SM00200; SEA; 1.
DR SMART; SM00202; SR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS01209; LDLRA_1; 2.
DR PROSITE; PS00068; LDLRA_2; 2.
DR PROSITE; PS00740; MAM_1; 1.
DR PROSITE; PS00060; MAM_2; 1.
DR PROSITE; PS50024; SEA; 1.
DR PROSITE; PS00420; SRCR_1; FALSE_NEG.
DR PROSITE; PS50287; SRCR_2; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Signal-anchor; Glycoprotein; Myristate; Hydrolase;
KW Serine protease; Zymogen; Transmembrane; Repeat.
FT CHAIN 52 117 NON-CATALYTIC M CHAIN (MINI CHAIN).
FT CHAIN 118 799 NON-CATALYTIC H CHAIN (HEAVY CHAIN).
FT CHAIN 800 1034 CATALYTIC L CHAIN (LIGHT CHAIN).
FT DOMAIN 1 18 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 19 47 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT DOMAIN 48 1034 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 52 169 SEA.
FT DOMAIN 197 238 LDL-RECEPTOR CLASS A 1.
FT DOMAIN 240 349 CUB 1.
FT DOMAIN 357 519 MAM.
FT DOMAIN 539 649 CUB 2.
FT DOMAIN 656 694 SRCR.
FT DOMAIN 800 1034 SERINE PROTEASE.
FT ACT_SITE 840 840 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 891 891 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 986 986 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT LIPID 2 2 MYRISTATE (POTENTIAL).
FT DISULFID 199 212 BY SIMILARITY.
FT DISULFID 206 225 BY SIMILARITY.
FT DISULFID 219 236 BY SIMILARITY.
FT DISULFID 658 670 BY SIMILARITY.
FT DISULFID 665 683 BY SIMILARITY.
FT DISULFID 677 692 BY SIMILARITY.
FT DISULFID 787 911 INTERCHAIN (BY SIMILARITY).
FT DISULFID 825 841 BY SIMILARITY.
FT DISULFID 925 992 BY SIMILARITY.
FT DISULFID 956 971 BY SIMILARITY.
FT DISULFID 982 1010 BY SIMILARITY.
```

[illegible][illegible]



Db 390 CLONRPAKELTVVLGDRHNQCEQCOTLAVRDYRLHEAFS---PITYQHDLALVRLQE 446  
Qy 155 AFQ--FGH---FVGPICLPE--LREQFEAGFICTTAGWGRLTGGVLSQVLQEVNLPILT 207  
Db 447 SADGCCAHPSPFVOPVCLPSTAARPAESEAACEVAGWHQFEGGEYSFLOEAQVPLID 506  
Qy 208 WEECVAAALLTKRPISGKTF----LCTGFPDGGRDACQDGGSLMCRNK--KGAWTLAG 261  
Db 507 PORCSAP-----DVHGAFTQGMCLCAGFLEGGTDACQDGGSLMCRNK--KGAWTLAG 261  
Qy 262 VTSWGLCGGRGNRRNVRKSDQGSFGIFTDISKVLSTIHEH 301  
Db 561 IVSWGSGCG----NRLK-----PGVYTDVANYLAWIREH 590

Search completed: March 25, 2003, 03:58:57  
Job time : 22 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 03:02:44 ; Search time 62 Seconds  
(without alignments)

1016.942 Million cell updates/sec

Title: US-09-735-713a-2

Perfect score: 1636

Sequence: 1 MSLKMLISRNKLLLLGIVF.....IFTDISKVLWIHEIQTGN 306

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_21.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mhc.\*

8: sp\_Organelle.\*

9: sp\_phase.\*

10: sp\_plant.\*

11: sp\_rodent.\*

12: sp\_virus.\*

13: sp\_vertebrate.\*

14: sp\_unclassified.\*

15: sp\_rvirus.\*

16: sp\_bacteriap.\*

17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	688.5	42.1	1004	13 P79953	P79953 xenopus lae
2	657	40.2	974	13 Q90WD8	Q90WD8 bufo japoni
3	626	38.3	1524	13 Q91674	Q91674 xenopus lae
4	496	30.3	277	5 Q96899	Q96899 scolopendra
5	486	29.7	787	5 Q9VEY6	Q9VEY6 drosophila
6	480.5	29.4	799	11 Q9DBI0	Q9DBI0 mus musculus
7	475.5	29.1	335	11 Q8VIF2	Q8VIF2 mus musculus
8	475.5	29.1	624	11 Q9DAT3	Q9DAT3 mus musculus
9	469.5	28.7	624	11 Q91Y47	Q91Y47 mus musculus
10	462.5	28.3	467	5 Q967X8	Q967X8 panulirus a
11	462	28.2	453	11 Q8VDE0	Q8VDE0 mus musculus
12	458	28.0	638	11 Q8E0P5	Q8E0P5 mus musculus
13	450	27.5	339	11 Q99L44	Q99L44 mus musculus
14	446	27.3	643	6 Q97506	Q97506 sus scrofa
15	440.5	26.9	433	13 Q90YK1	Q90YK1 brachydanio
16	436.5	26.7	317	13 Q9DGR3	Q9DGR3 xenopus lae

17	435.5	26.6	310	11 Q91XC4	Q91XC4 mus musculus
18	430.5	26.3	310	11 Q9QY29	Q9QY29 mus musculus
19	430	26.3	767	13 Q9DGR2	Q9DGR2 xenopus lae
20	428.5	26.2	581	5 Q96015	Q96015 drosophila
21	428.5	26.1	1047	5 Q9VZH2	Q9VZH2 drosophila
22	427.5	26.1	387	5 Q9XY57	Q9XY57 ctenocephal
23	426	26.0	314	5 Q9VR15	Q9VR15 drosophila
24	426	26.0	624	6 Q95ME7	Q95ME7 oryctolagus
25	423.5	25.9	325	5 Q15944	Q15944 sarcophaga
26	423.5	25.9	616	6 Q97507	Q97507 sus scrofa
27	423	25.9	256	5 Q9XYI1	Q9XYI1 rhyzopertha
28	422.5	25.8	267	5 Q9BK47	Q9BK47 luidia foli
29	422.5	25.8	456	6 Q9TTR0	Q9TTR0 canis famil
30	421	25.7	273	6 Q9XSM1	Q9XSM1 ovis aries
31	421	25.7	855	11 Q91J17	Q91J17 rattus norv
32	420.5	25.7	321	4 Q96R28	Q96R28 homo sapien
33	420.5	25.7	1047	5 Q24019	Q24019 drosophila
34	418.5	25.6	329	6 Q9GL10	Q9GL10 ovis aries
35	417	25.5	264	11 Q9D7P8	Q9D7P8 mus musculus
36	417	25.5	264	11 Q9ER05	Q9ER05 mus musculus
37	414.5	25.3	300	4 Q96EF3	Q96EF3 homo sapien
38	414.5	25.3	573	5 Q9V516	Q9V516 drosophila
39	414	25.3	254	5 Q8T637	Q8T637 aedes aegypt
40	414	25.3	264	11 Q9EQ28	Q9EQ28 rattus norv
41	414	25.3	268	5 Q46151	Q46151 pacifastacu
42	413	25.2	263	11 Q9DC86	Q9DC86 mus musculus
43	413	25.2	405	4 Q96E86	Q96E86 homo sapien
44	412	25.2	263	11 Q9CR35	Q9CR35 mus musculus
45	409.5	25.0	468	5 Q9U0G3	Q9U0G3 pacifastacu

## ALIGNMENTS

### RESULT 1

P79953	ID	P79953	PRELIMINARY;	PRT;	1004 AA.
AC	P79953;				
DT	01-MAY-1997	(TRENBLREL. 03, Created)			
DT	01-MAY-1997	(TRENBLREL. 03, Last sequence update)			
DT	01-JUN-2002	(TRENBLREL. 21, Last annotation update)			
DE	Oviductin.				
OS	Xenopus laevis (African clawed frog).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;				
OC	Xenopodinae; Xenopus.				
OX	NCBI_taxid=8335;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE-99184825; PubMed-10084976;				
RA	Lindsay L.L., Wieduwilt M.J., Hedrick J.L.;				
RT	"Oviductin, the Xenopus laevis oviductal protease that processes egg				
RT	envelope glycoprotein gp43, increases sperm binding to envelopes, and				
RT	is translated as part of an unusual mosaic protein composed of two				
RT	protease and several CUB domains.";				
RL	Biol. Reprod. 60:989-995(1999).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RA	Yang J.C., Lindsay L.L., Hedrick J.L.;				
RL	Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.				
CC	-I- SIMILARITY: CONTAINS 3 CUB DOMAINS.				
CC	-I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE				
CC	TRYPSIN FAMILY				
DR	EMBL; U81291; AAB53972.1; -				
DR	HSSP; P00763; IDPO.				
DR	MEROPS; S01.240; -				
DR	InterPro: IPR001314; Chymotrypsin.				
DR	InterPro: IPR000859; CUB domain.				
DR	InterPro: IPR001254; Ser_protease_Try.				
DR	Pfam; PF00431; CUB; 2.				
DR	Pfam; PF00089; trypsin; 2.				
DR	PRINTS; PR00722; CHYMOTRYPSIN.				
DR	SMART; SM00042; CUB; 3.				



```
Qy 24 GKSAALSLPAPSCGQSLVKVQPNWYNIFSRILGSGQVEKGYSPWQVSLKROKHICGG 83
Db 32 GETAELCKGTPEIGD-----EPDLEF--SRIVGGGDAVGGQPTVSLKLNHRHICGG 84
Qy 84 SIVSQWVITAAHCII--ANRNIVSTLNVATAGEYDLSTQDPEQTLTIETVTHPHFSTKK 141
Db 85 SIVRKDMVVTAAHCVTPVTEIKSHMTVIVGEYDQVMDSOEQSIPVSHIEPHPNRYGDG 144
Qy 142 PMDYDIALLKAGAFQCHFGVPTCLPELREQFAGFCTTAGHGRLTGEGVGLSQVLOEV 201
Db 145 NMGDIALVFLVSKPIIFGSGQVQICLPQVGEKIEAGTLCVSSGWRLEENGDLSPLVLOEV 204
Qy 202 NLPILTWEECAALTLTKRPISGTFCTGPPDGGRCACQDQSGSLMCRNKKGAWTLAG 261
Db 205 KLPVVDNCTCHAVLEIGHVPLDDTMLCAGEPEGMDACQDQSGGPFVCRRRSGVWFLAG 264
Qy 262 VTSNGLGCGRW--RNNVRKSDQSGPGFTIDISKVLSWTHIEHQG 305
Db 265 CVSMGLGCGRWGAKIIRISGSGPAIFSRVSSVLDLRLPPKLTG 309

RESULT 4
O96899 ID O96899 PRELIMINARY; PRT; 277 AA.
AC O96899;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE Plasmidogen activator spa.
OS Scolopendra subspinosipes.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Myriapoda; Chilopoda;
OC Pleurostigmophora; Scolopendromorpha; Scolopendridae; Scolopendra.
OX NCBI_TaxID=55038;
RN [1]
RP SONG Y.-D., You W.K., Kim K.Y., Chung K.H., Park D.-H.;
RT "Cloning of a cDNA for a novel plasminogen activator from a Korean
RL centipede Scolopendra.";
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; J079521; AAD00320.1; -.
DR HSP; P00763; IDPO.
DR MEROPS; S01.122; -.
DR InterPro; IPR001314; Chymotrypsin.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYD-SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 277 AA; 30172 MW; EC62F80C8CB246C CRC64;

Query Match 30.38; Score 496; DB 5; Length 277;
Best Local Similarity 38.18; Pred. No. 4e-39;
Matches 112; Conservative 46; Mismatches 98; Indels 38; Gaps 12;

Qy 10 NKLILLGIVFERCKSAALSIPRAPSGQSLVKVQPNWYNIFSRILGSGQVEKGYSPW 69
Db 2 NSFTILI-VTF-----SLAFSGRCG---IKNGP--MLDEFNRIVGGEAEPGEPW 47
Qy 70 QVSLK----QRKHICGSGIVSPQVITAAHCIAANRNIVSTLNVATAGEYDLSTQDPEQGT 125
Db 48 QISLQVSWYSGYHYCGSGILDESNTVTAACHVEGMN-PSDLRLILAGEHNKKEDGTQW 106
Qy 126 LTETVTHPHFSTKKPMYDIALLKAGAFQF-GHFVGPICLPFLREQFAGFTCTTAG 184
Db 107 QVDVIDIIMHKDY-VYSTLENDIALKLAEPDLTPTAVGSCICLPSQNNQFSGH-CIVTG 164
Qy 185 WGRLTGEGVLSQVLOEVNLPILTWECAVALLTLKRPISGKTFCTGPPDGGRCACQDGS 244
```

```
Db 165 WGSYREGGNSPNILQKVSPLMTDEECSEYNIIV-----DTMLCAGYAEGKDACQDGS 218
Qy 245 GGSMLCRNKKCAWTLAGVTSMGLCGGRWNRNVRKSDQSGPGIFTDTSKVLWSWI 298
Db 219 GGPLVCPNGDGTYSLAGIVSMGICGAQP-RN-----PGVYTVQVSKFLDWI 262

RESULT 5
O9VEY6 ID O9VEY6 PRELIMINARY; PRT; 787 AA.
AC O9VEY6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE SB gene product.
GN SB OR CG4316.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananatiades P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaiswal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.M.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AE003712; AF555277.1; -.
DR HSP; P00763; IDPO.
DR MEROPS; S01.225; -.
DR FlyBase; FBgn0003319; Sb.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR002965; P_rich_extensn.
DR InterPro; IPR001254; Ser_protease_Try.
```

CC	TRYPSIN FAMILY.		
DR	EMBL; AK004939; BAB23684.1; -.		
DR	HSP; P00763; IDPO.		
DR	MEROPS; S01.308; -.		
DR	MGD; MGI:1919003; 1300008A22rik.		
DR	InterPro; IPR001314; Chymotrypsin.		
DR	InterPro; IPR000859; CUB_domain.		
DR	InterPro; IPR002172; LDL_recept_A.		
DR	InterPro; IPR001254; Ser_protease_Try.		
DR	Pfam; PF00057; ldl_recept_a; 2.		
DR	Pfam; PF00089; trypsin; 1.		
DR	PRINTS; PR00722; CHYMOTRYPSIN.		
DR	PRINTS; PR00261; LDLRECEPTOR.		
DR	SMART; SM00042; CUB; 1.		
DR	SMART; SM00192; Ldla; 3.		
DR	SMART; SM00020; Tryp_SPC; 1.		
DR	PROSITE; PS01180; CUB; 1.		
DR	PROSITE; PS01209; LDLRA1; 1.		
DR	PROSITE; PS50068; LDLRA2; 3.		
DR	PROSITE; PS50240; TRYPSIN_DOM; 1.		
DR	PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.		
DR	PROSITE; PS00135; TRYPSIN_SER; 1.		
SW	Glycoprotein; Hydrolase; Serine protease.		
KQ	SEQUENCE 799 AA; 89557 MW; 16315A646A4D5288 CRC64;		
QY	Query Match	29.4%; Score 480.5; DB 11; Length 799;	
DB	Best Local Similarity	38.1%; Pred. No. 4.8e-37;	
DB	Matches	96; Conservative	45; Mismatches 88; Indels 23; Gaps
QY	54	SRILGGQVBKSGYPMQVSLKQRKHICGSGIVSPQWVITAAHCIAINRNVIS--TLNVTA 111	
DB	563	SRIVGGTVSSGESEWPMQASLIQRGRHICGGALIDRNVITAAHCFQEDSMASPKLMTVFL 6222	
QY	112	GEYDLSQTDQEGTTLTIETVLIHPHSTKKPMQYDIALKWKAGAFQGFHFVGPICLPELR 171	
DB	623	GKMRQNRWPEGVSEFKVSRFLHP-VHEEDSHYDVALQLDHPVIVYSATVRPVCULPARS 681	
QY	172	EQFEAGICTAGNRLTEGGVLSQVLQENPLPILTWEECAALLTLKRPISGKTFLCGT 231	
DB	682	HFEFGQHCHITWGAGQREGGVPSTLQKVDVLVPQDLCSEAYRQVSP----RMLCAG 737	
QY	232	FPDGRDACQDGGSLMCRNKKGAWTLAGVTSWGLCGCGRWNNVYKSDQSGP---GIF 288	
DB	738	YRKGKDACQDGGSLVCREPSGRMFLAGLSVWGLGCGR-----PNFFGVY 784	
QY	289	TDISKVLWSIHE 300	
DB	785	TRVTRVINWIIQ 796	
RESULT 7			
Q8VIF2	ID	Q8VIF2	PRELIMINARY; PRT; 335 AA.
DT	OC	Q8VIF2;	
DT	DT	01-MAR-2002 (TrEMBLrel. 20, Created)	
DT	DT	01-MAR-2002 (TrEMBLrel. 20, Last sequence update)	
DT	DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)	
DE	DE	Testis serine protease2.	
GN	GN	TESSP2.	
OC	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
ON	NCBI_TaxID=10090;		
RX	[1]		
RP	SEQUENCE FROM N.A.		
RA	Matsui H., Takano N., Takahashi T.;		
RT	"nresp2,serine protease specifically expressed in mouse testis.";		
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AB052292; BAB78735.1; -.		
DR	InterPro; IPR001314; Chymotrypsin.		
DR	InterPro; IPR001254; Ser_protease_Try.		
DR	Pfam; PF00089; trypsin; 1.		
DR	PRINTS; PR00722; CHYMOTRYPSIN.		

```
DR SMART; SM00020; Tryp_Spc; 1.
DR PROSITE; PS0240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
DR Protease.
KW SEQUENCE. 335 AA; 36682 MW; E8FC667DC9427EFC CRC64;

Query Match          29.1%; Score 475.5; DB 11; Length 335;
Best Local Similarity 37.1%; Pred. No. 4.6e-37;
Matches 102; Conservative 49; Mismatches 89; Indels 35; Gaps 9;

Qy 33 KAPSCGSLVKVQPNVFNIFSRILGSGQVEKGYPMQVSLKQKHKHGGSIYSPQWVI 92
Db 62 RSPFNFSLVGQCP-----FMKIMGVDAEKGKWPQVSVVRHMHVCGGSLNSQWVL 115
Qy 93 TAAHCANRNIVSLNVTAGYDLSQTDPGEQTLTIETVTHPHFSTYKPKMDYDIALKLM 152
Db 116 TAAHCIIYSR-----IQYNKVGDRSVYRONT-SLVIPIKTIHFVHPKFTIVYVKNIDIALKL 171
Qy 153 AGAFQFGHFVGPICLPRLREQFEAGFICTTAGWGLRLEGG--VLSQVLQELNLPILTWE 210
Db 172 QHPVNFNTNIVPCIPSEFPVKAGTKCWVTGCKLPGAPDVPTEILQEVQNVILYEE 231
Qy 211 CVAALLTLKRPIG-----KTFLETGFPDGGGRDACQDGSGLMC--RNKKGAWTLAGVT 263
Db 232 CNE---MLKKATSSVDLVKRGWVCGYKERGKDACQDGSGLMSEFENK---NWQGVV 285
Qy 264 SWGLGCGRGWRNNVRKSDQSGPGIFTDISKVLWSI 298
Db 286 SWGISCCR-----KGYPGVYTDVAFYSKWL 310

RESULT 8
Q9DAT3 PRELIMINARY; PRT; 624 AA.
AC Q9DAT3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE 1600027G01Rk protein (RIKEN cDNA 1600027G01 gene).
GN 1600027G01Rk.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PLACENTA;
RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Azawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Anon H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Strausberg R.;
```

```
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; AK005546; BAB24114.1; -.
DR EMBL; BC019485; AAH19485.1; -.
DR HSP; P00750; IRTF.
DR MEROPS; S01.213; -.
DR MGD; MGI:1919281; 1600027G01Rk.
DR InterPro; IPR000177; Apple.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR003014; PAN.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00024; PAN; 4.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00005; APPLIEDOMAIN.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00223; APPLE; 4.
DR SMART; SM00202; Tryp_Spc; 1.
DR PROSITE; PS00495; APPLE; UNKNOWN_3.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 624 AA; 69788 MW; 0BEDDEBC56009E97 CRC64;

Query Match          29.1%; Score 475.5; DB 11; Length 624;
Best Local Similarity 35.1%; Pred. No. 1e-36;
Matches 99; Conservative 47; Mismatches 113; Indels 23; Gaps 6;

Qy 23 RGKSAALSRLKAPSCGSGSLVKVQPNVFNIFSRILGSGQVEKGYPMQVSLKQKHKICG 82
Db 365 RGGISGYSLRLCKMDNVCTTKINP-----RVVGAASVHGEWPMQVTLHISQGLCG 416
Qy 83 GSIVSPQWVITAAHCANRNIVSLNVTAGYDLSQTDPGEQTLTIETVTHPHFSTKPK 142
Db 417 GSIIQNWILTAACHFCGSIETPKLRYGVGVNQSEINEGTAFFRVOEMIIHDQVTTAE- 475
Qy 143 MDYDIALKMAAGAFQGHFVGPICLPRLREQFEAGFICTTAGWGLRLEGGVLSQVLQEVN 202
Db 476 SGYDIALKLESAMNYTDFQRPICLPKSGDRNAVHTCEWGTGVTALRGVEQSTLQAK 535
Qy 203 LPILTWECEVAALLTLKRPIGKTFCTGFPDGGGRDACQDGSGLMCRNKGAWTLAGV 262
Db 536 VPLVSNCEQTRY--RRHKITNK-MICAGYKEGKDTCKDGGPLSCK-YNGVWHLVGI 591
Qy 263 TSWGLGCGRGWRNNVRKSDQSGPGIFTDISKVLWSIHEHIQT 304
Db 592 TSWGEGCGQKER-----PGVYTNVAKYVDWILEKTQT 623

RESULT 9
Q91Y47 PRELIMINARY; PRT; 624 AA.
ID Q91Y47;
AC Q91Y47;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Coagulation factor XI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=B10.WR; TISSUE=LIVER;
RX MEDLINE=97385041; Pubmed=9242536;
RA Gallani D., Sun M.F., Sun Y.;
RT "A comparison of murine and human factor XI.";
RL Blood 90:1055-1064(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=B10.WR; TISSUE=LIVER;
RA Gallani D., Sun M.F., Sun Y.;
```

Best Local Similarity 38.2%; Pred. No. 1.2e-35;  
Matches 105; Conservative 40; Mismatches 97; Indels 33; Gaps 9;

QY 35 PSCGOSLVKVPWNYFNIFSRILGSGVEKSYPKWYSLKROKH-ICGGSIIVPQWVIT 93  
| | | : | | : | | | | : | | | | | | : | | | | | | : | | | |  
Db 216 PSCAGNV-----NRATRVGGFEVNEVYMQVLLVTRDMVICGGSIISSQWL 267  
| | | : | | : | | | | : | | | | | | : | | | | | | : | | | |

QY 94 AAHCIANRNIYSTLNVTAGEYDLSOTDPGEQTLETIETV--TIHFFSTKKPMDYDIALLK 151  
| | | : | | : | | : | | : | | : | | : | | : | | : | | : | | | |  
Db 266 AACHVDGNIGVVL---VGDNHFASDTTTSRLVEVVQLISHPDYDS-STVNDMALLR 323  
| | | : | | : | | : | | : | | : | | : | | : | | : | | : | | | |

QY 152 MAGAFQGHHFVGPICLPPELRQFPAQFTCTTAGNRLTECGVLSQLQEVNLPLTWEEC 211  
| | : | | : | | | : | | | | | | | | | | | | | | | | | | : | | | |  
Db 324 LGEALFRETAVAPCLPSNPTEADYGATVATVTWGAAATEGGSMSTLQEDVPDLTTAAC 383  
| | | : | | : | | | | : | | | | | | | | | | | | | | | | | : | | | |

QY 212 VAULLTKRPISGKTFLCTGFPGDGRCACGDSCGLMCRNKKKAWTLAGVTSWGLCCGR 271  
| | : | | : | | | : | | | | | | | | | | | | | | | | | | : | | | |  
Db 384 SSWSYSLT-----ANMMCAGSNEGKSCQDGGSP-MVVSATSNYEIQGVVSWGRCAR 437  
| | | : | | : | | : | | : | | : | | : | | : | | : | | : | | | |

QY 272 GWRRNVRSKDQGPSGIPTDISKLVSWIHEHIQTGN 306  
| | | : | | : | | : | | : | | : | | : | | : | | : | | : | | | |  
Db 438 -----FGFCGYVARVTEYLEWTAAN--TGN 460  
| | | : | | : | | : | | : | | : | | : | | : | | : | | : | | | |

RESULT 11

ID Q8VDEO PRELIMINARY; PRT; 453 AA.  
AC Q8VDEO;  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE TMPRSS3 protein.  
GN TMPRSS3.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
NCBI\_TaxID=10090;  
[1]  
SEQUENCE FROM N.A.

RA Guipponi M., Scamuffa N., Scott H.S., Rossier C., Antonarakis S.E.;  
RT "Isolation and characterization of the mouse tmprss3 gene."  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ300738; CAC83350.1; -  
MGD; MGI:2155445; Tmprss3.  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR002172; LDL\_recept\_A.  
DR InterPro; IPR001254; Ser\_protease\_Try.  
DR Pfam; PF00057; ldl\_recept\_a; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPSIN.  
DR SMART; SM00192; LDla; 1.  
DR SMART; SM00202; SR; 1.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS01209; LDLRA\_1; UNKNOWN\_1.  
DR PROSITE; PS50068; LDLRA\_2; 1.  
DR PROSITE; PS50287; SRCR\_2; 1.  
DR PROSITE; PS50240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
DR PROSITE; PS00135; TRYPSIN\_SER; UNKNOWN\_1.  
SQ SEQUENCE 453 AA; 49529 MW; 21E5697DC8781BD3 CRC64;

Query Match 28.2%; Score 462; DB 11; Length 453;  
Best Local Similarity 33.1%; Pred. No. 1.3e-35;  
Matches 102; Conservative 57; Mismatches 115; Indels 34; Gaps 8;

QY 1 MSCLKLIISRKNLLLGIVFPFER-KSAALSPLRAPSCGOSLVKVPWNFIIFSRTLGG 59  
| : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | | |  
Db 170 VSHILLSDDKVTAHLHSVYNRECTGHVTVTLKCSACGTFT-----GYSPIRVGG 220  
| : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | | |

QY 60 SQVEKGSPQWQSUKQRKHICGGSIVSPQWVIITAASHCIANRRNIVSTLNTAYEDLSQT 119  
| : | | | | : | | | | | | | | | | | | | | | | | | | | | | : | | | |  
Db 221 NMSSLTOWPWQVWSLOFOGYHLHGCSIITPTFWIVTAAAHCVDVLDVHPKSWTVQVGI--VSIM 278  
| : | | | | : | | | | | | | | | | | | | | | | | | | | | | : | | | |





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 25, 2003, 01:17:33 ; Search time 1806 seconds  
(without alignments)  
2744.086 Million cell updates/sec

Title: US-09-735-713A-2  
Perfect score: 1636  
Sequence: 1 MSUKMLISRNKLLILLGIVF.....IFTDISKVLSTWIEHQTGN 306

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-Q/cgcn2\_1/USPTO\_spool/US09735713/runat\_18032003\_124142\_21598/app\_query.fasta\_1.455  
-DB=EST -QPMT=fastap -SUFFIX=fst -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09735713 -CGCN\_1.1.1906 -runat\_18032003\_124142\_21598 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_man:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	968	59.2	670	10	BB625475	BB625475 Mus muscu
2	480.5	29.4	3030	11	AK004939	AK004939 Mus muscu
3	475.5	29.1	2246	11	AK005546	AK005546 Mus muscu
4	426	26.0	3248	11	BC008514	BC008514 Mus muscu
5	424.5	25.9	1052	13	BI554641	BI554641 603235988
6	422.5	25.8	999	9	AL551470	AL551470 AL551470
7	420	25.7	1188	11	BC034294	BC034294 Homo sapi
8	418.5	25.6	790	13	BJ509337	BJ509337 BJ509337
9	417	25.5	868	11	AK009019	AK009019 Mus muscu
10	415.5	25.4	1056	14	BM919200	BM919200 AGENCOURT
11	415.5	25.4	1323	11	AK014645	AK014645 Mus muscu
12	414	25.3	1629	11	AK010640	AK010640 Mus muscu
13	413	25.2	885	11	AK003060	AK003060 Mus muscu
14	412.5	25.2	921	13	BI833975	BI833975 603085088
15	412	25.2	876	11	AK007765	AK007765 Mus muscu
16	412	25.2	877	11	AK003079	AK003079 Mus muscu
17	412	25.2	877	11	AK007815	AK007815 Mus muscu
18	412	25.2	877	11	AK008729	AK008729 Mus muscu
19	412	25.2	877	11	AK008927	AK008927 Mus muscu
20	412	25.2	879	11	AK008888	AK008888 Mus muscu
21	412	25.2	943	14	BQ222778	BQ222778 AGENCOURT
22	411.5	25.2	807	13	BI834199	BI834199 603084101
23	410	25.1	870	11	AK007333	AK007333 Mus muscu
24	407	24.9	1050	11	AK006271	AK006271 Mus muscu
25	406	24.8	876	11	AK007566	AK007566 Mus muscu
26	405.5	24.8	905	14	BQ729161	BQ729161 AGENCOURT
27	403	24.6	870	14	BM919892	BM919892 AGENCOURT
28	400.5	24.5	855	14	BQ956801	BQ956801 AGENCOURT
29	399.5	24.4	671	13	BI771604	BI771604 603058729
30	399	24.4	890	11	AK008644	AK008644 Mus muscu
31	399	24.4	1085	14	BM919045	BM919045 AGENCOURT
32	398.5	24.4	780	13	BI838552	BI838552 603086213
33	396.5	24.2	1115	14	BM918560	BM918560 AGENCOURT
34	393.5	24.1	876	9	AL555870	AL555870 AL555870
35	391.5	23.9	1024	11	AK006746	AK006746 Mus muscu
36	390.5	23.9	936	9	AL578261	AL578261 AL578261
37	390.5	23.9	1814	11	AK002694	AK002694 Mus muscu
38	390	23.8	896	11	AK009129	AK009129 Mus muscu
39	390	23.8	897	11	AK010149	AK010149 Mus muscu
40	387	23.7	668	13	BM641281	BM641281 170006873
41	386	23.6	695	13	BM576074	BM576074 170006871
42	386	23.6	838	13	BI834360	BI834360 603084395
43	386	23.6	925	14	BQ228286	BQ228286 AGENCOURT
44	385.5	23.6	870	11	AK007773	AK007773 Mus muscu
45	385.5	23.6	1006	11	AK008460	AK008460 Mus muscu

ALIGNMENTS

RESULT 1  
BB625475  
LOCUS BB625475 RIKEN full-length enriched, adult male epididymis Mus  
DEFINITION BB625475 musculus CDNA clone 9230106D23 5', mRNA sequence.  
ACCESSION BB625475  
VERSION BB625475.1 GI:16463737  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 670)  
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda







and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence[5' GAGAGAGAGCGCGCAATTAATCTCGAGTAAATTAATTAATCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end:XhoI; 3' end: SstI. Host: SOLR.

```

FEATURES
source
    Location/Qualifiers
        1..2246
            /organism="Mus musculus"
            /strain="C57BL/6J"
            /db_xref="FANTOM_DB:1600027G01"
            /db_xref="MGD:MGI:1904755"
            /db_xref="taxon:10090"
            /clone="1600027G01"
            /sex="female"
            /tissue_type="placenta"
            /clone_lib="RIKEN full-length enriched mouse cDNA library"
            /dev_stage="adult"
        94..1968
            /note="data source:SPTR, source key:P03951, evidence:ISS
            homolog to COAGULATION FACTOR XI PRECURSOR (EC 3.4.21.27)
            (PLASMA THROMBOPLASTIN ANTECEDENT) (PTA)
            putative"
CDS

```

BASE COUNT	666 a	509 c	536 g	535 t
ORIGIN				

Alignment Scores:			
Pred. No.:	5,11e-42	Length:	2246
Score:	475.50	Matches:	99
Percent Similarity:	51.77%	Conservative:	47
Best Local Similarity:	35.11%	Mismatches:	113
Query Match:	29.06%	Indels:	23
DB:	11	Gaps:	6
US-09-735-713A-2 (1-306) x AK005546 (1-2246)			
Qy	23	ArgGlyLysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGlnSerLeuVal	42
		. . . :	
Db	1186	AGGGGAGGCATCTCTGGATACACTGAGCGGTGTGCAAAATGGATAATGTGTGCACAAC	1245
Qy	43	LysValGlnProTrpAsnTyrPheAsnIlePheSerArgIleLeuGlyGlySerGlnVal	62
		. . . :	
Db	1246	AAATCAACCC-----AGAGTGTAGGAGGCTGCCTCT	1281
Qy	63	GlulysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIleCysGly	82
		:   :     :	
Db	1282	GTTACGGTGAGTGGCCATGGCAGGTGACTCTGCACATCAGCCAGGCACACCTGTGTGGA	1341
Qy	83	GlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsnArgAsn	102
		:     :     :	
Db	1342	GGCTCATCTATTTGGAAACCAATGGATCTGACGAGCATCTATTCTTCTCTGGGATAGAG	1401
Qy	103	IleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAspProGly	122
		:	
Db	1402	ACACCTAAAAGCTGGCTCTACCGTGGCATTTGTAATCAATCAGAAATAATGAAGG	1461
Qy	123	GlucGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLysLysPro	142
		:   : :   :	

FEATURES

BASE COUNT  
ORIGIN

Pred. No.:	3.14e-36	Length:	3248
Score:	426.00	Matches:	90
Percent Similarity:	51.33%	Conservative:	45
Best Local Similarity:	34.22%	Mismatches:	90
Query Match:	26.04%	Indels:	38
DB:	11	Gaps:	7

QY 54 SerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu 73

Db 2009 GCTCGCGTGGTGGCCACGAATCCGACGAGGCGAGTGGCCCTGGCAGGTGAGCCTC 2068

[illegible][illegible]

Qy 113 GluTyrAspLeu-----SerGlnThrAspPro 121

50 2174 GACACACGATGATGACGGCCATCCGAGGATCGACCAAGCGCAGTGGCTCT 2233

[illegible]

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

Db 2351 GTGCGCCCCATCTGCCCTGCCCTGATGCTACCCATGTCTTCCCTGCTGGCAAGGCCATCTGG 2410

Ov 182 ThrAlaGlyTTrpGlyArgLeuThrGlyGlyValLeuSerGlyValLeuGlyVal 201

Db 2411 GTCACAGGCTGGGGGCACACAAAGAGGGAGGTACCGAGCGTGATCTGCAGAGGGT 2470

[illegible][illegible]

Db 2531 CGA-----ATGATGTGTGGGTTTCTCAGTGGGGTGTGGAC 2569

QY 239 AIAcYSGLINGLYAspserGlyGlyserLeumetCysArgasnLysLysGlyAlaTrpThr 258  
... ||||| ||||| |||||

— — — — —

Db	2630	CAGCTGGTGTGGTGAAGCTGGGGTGAAGCTGCG	-----	2662
Qy	279	LysSerAspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIle	298	
Db	2663	--GCTCAGAGGACAGCCAGCGGTGTACACAGGCTCCTGTAGTTCGGGACTGGATC	2719	
Qy	299	HisGluHis	301	
Db	2720	AAAGAGCAC	2728	
RESULT	5			
BI554641				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				
FEATURES				
source				
BASE				
ORIGIN				
Alignment				
Scores:				
Score:				
Percent Similarity:				
Best Local Similarity:				
Query Match:				
DB:				
US-09-735-713A-2				
Qy	54	SerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu	73	
Db	68	AGCGTATTGTGGCGGAGCCGTGTCTCCGAGGGTGAGTGGCCATGGCAGCCAGCCCTC	127	
Qy	74	LysLsnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpValIleThr	93	
Db	128	CAGATTCCGGGGTGCACACATCTGTGGGGGGGCTCTCATCGCTGACCGCTGGGTCAACG	187	
Qy	94	AlaAlaHisCysIleLeuAlaAsnArgAsnIleValSer-----ThrLeuAsnValThrAla	111	
Db	188	GCCGCCACACGTGTTCCAGGAGGACACAGATCGGCTCCCGGAAGCTGTGGACCGTGTCCCTG	247	





TITLE Medaka EST Project in Takeda's lab  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Tadasu Shin-i  
 Center For Genetic Resource Information  
 National Institute of Genetics  
 1111 Yata, Mishima, Shizuoka 411-8540, Japan  
 Tel: 81-559-81-6856  
 Fax: 81-559-81-6855  
 Email: tshin@genes.nig.ac.jp.

FEATURES Location/Qualifiers

source

1. .790

/organism="Oryzias latipes"

/strain="d-r"

/db\_xref="taxon:8090"

/clone="MF01FSA007022"

/clone\_lib="MF01FSA cdna"

/sex="mixture of female and male"

/tissue\_type="whole embryo"

/dev\_stage="fry stage 40"

BASE COUNT 176 a 204 c 231 g 178 t 1 others

ORIGIN

Alignment Scores:  
 Pred. No.: 2.66e-36 Length: 790  
 Score: 418.50 Matches: 93  
 Percent Similarity: 53.88% Conservative: 46  
 Best Local Similarity: 36.05% Mismatches: 84  
 Query Match: 25.58% Indels: 35  
 DB: 13 Gaps: 10

US-09-735-713a-2 (1-306) x BJ509337 (1-790)

QY 54 SerArgLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu 73  
 DB 788 ACCAGGTGTCGGTGAAGATGTCGAGCGCACAGCTGGCCCTGGCAGGTGTCCTG 729  
 QY 74 Lys-----GlnArgGlnLysHisIleCysGlyGlySerIleValSerProGln 89  
 DB 728 CAGTACTGAGTGTAACACTTATTACCACACCTGTGGGGTACCTGATCTCCACGAG 669  
 QY 90 TrpValIleThrAlaAlaHisCysIleAlaAsnArgAsnIleValSerThrLeuAsnVal 109  
 DB 668 TGGGTCTCTACTGCTGCTACTGCTCATGCGAAGTCG-----ACGTACAGAGTG 621  
 QY 110 ThrAlaGlyGluTyrAspLeuSer---GlnThrAspProGlyGluGlnThrLeuThrIle 128  
 DB 620 TACATCGGAACACGACCTGAGCGCCACACAGAGCTGGTCCATCGCCATCGAGCC 561  
 QY 129 GluThrValIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAla 148  
 DB 560 CAAAGGATCATCTCCACGAGAACTGGGACTCTTACAAC--ATCAGAAATGACATTGCC 504  
 QY 149 LeuLeuLysMetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuPro 168  
 DB 503 CTGATCAAACTGTCAACTCCAGTCACCTTTTCCAAACACCATTTGGGCTGCTGCTCCA 444  
 QY 169 GluLeuArgGlnGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeu 188  
 DB 443 AAGCTGGTGAATCTCGGATGACGGTGCCTCTGATGTACCGGCTGGGACGTCCT 384  
 QY 189 ThrGluGlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTrp 208  
 DB 383 TGGACTGGAGTCCCATCTGATATCTCTGACAGAGCCCTCTCCAGTGGTGGACAC 324  
 QY 209 GluGluCysValAlaAlaLeuLeuThrLeuLysArgPro-----IleSer 223  
 DB 323 TCCACCTGC-----AGCAGGCTGACTGCTGGTGGGCACTATCTGTC 285  
 QY 224 GlyLysThrPheLeuCysThrGlyPheProAspGlyArgaspAlaCysGlnGlyAsp 243  
 DB 284 ACCACCAACATGATTTGTCGCGGA---GGAGCGGAGAGTGGCTAGCTCAACGGAGAC 228  
 QY 244 SerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAlaGlyValThr 263

Db 227 TCTGGTGGCCCTGGAACCTGTCGCAACCCCTGATGGCTCTGGGATGTCACCGAATTGTG 168  
 QY 264 SerTrpGly-----LeuGlyCysGlyArgGlyTrpArgAsnValArgLysSerAsp 281  
 Db 167 AGCTTCGGCTCCAGCCCTGGGTGC-----AACTACCCCAAGAAG----- 129  
 QY 282 GlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHis 299  
 Db 128 -----CCCTCTGCTTCACCAGAGTCAGCGGCTACATCAGCTGAGTCAAC 84  
 RESULT 9  
 AK009019 868 bp mRNA linear HTC 19-JAN-2002  
 LOCUS Mus musculus adult male tongue cDNA, RIKEN full-length enriched  
 DEFINITION library, clone:2300003F21:chymotrypsin-like, full insert sequence.  
 ACCESSION AK009019  
 VERSION AK009019.1 GI:12843558  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (strain:C57BL/6J) adult male tongue cDNA to mRNA,  
 clone\_lib:RIKEN full-length enriched mouse cDNA library  
 clone:2300003F21.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1  
 AUTHORS Carninci,P. and Hayashizaki,Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636  
 REFERENCE 2  
 AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,  
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20493374  
 PUBMED 11042159  
 REFERENCE 3  
 AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,  
 Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,  
 Sumi.N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,  
 Yamamoto,R., Matsumoto,H., Sakauchi,S., Ikegami,T., Kashiwagi,K.,  
 Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,  
 Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,  
 Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer  
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE 20530913  
 PUBMED 11076861  
 REFERENCE 4  
 AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,  
 Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,  
 Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,  
 Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,  
 Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,  
 Fletschmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,  
 Kuehl,P., Lewis,S., Matsuo,Y., Nikaudo,I., Pesole,G.,  
 Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,  
 Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,  
 Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,  
 Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,  
 Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,  
 Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,  
 Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,  
 Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,  
 Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,  
 Toyooka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,  
 Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.  
 and Hayashizaki,Y.  
 TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS

Nature 409 (6821), 685-690 (2001)  
21085660  
11217851  
5 (bases 1 to 868)  
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,  
Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,  
Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T.,  
Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,  
Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,  
Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,  
Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,  
Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,  
Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,  
Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I.,  
Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and  
Hayashizaki, Y.

TITLE  
JOURNAL

Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp,  
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)

## COMMENT

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for  
further details.  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues. First strand cDNA was primed with a primer  
[5'-GAGAGAGAAGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was  
prepared by using trehalose thermo-activated reverse transcriptase  
and subsequently enriched for full-length by cap-trapper. Second  
strand cDNA was prepared with the primer adapter of sequence [5'-  
GAGAGAGATTCTCGATTGTAATAATATCCCGCCCCC 3']. cDNA was cleaved  
with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.  
Host: SOLR.

## FEATURES

## source

Location/Qualifiers  
1..868  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="FANTOM\_DB:2300003F21"  
/db\_xref="MGD:MG1:1893096"  
/db\_xref="taxon:10090"  
/clone="2300003F21"  
/sex="male"  
/tissue\_type="tongue"  
/clone\_lib="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="adult"  
1..868  
/gene="Ctrl"  
19..813  
/gene="Ctrl"  
/note="chymotrypsin-like  
data source:MGD, source key:MG1:88558, evidence:ISS  
putative"  
/codon\_start=1  
/protein\_id="BAB26029.1"  
/db\_xref="GI:12843559"  
/db\_xref="MGD:MG1:88558"  
/translation="MLLSLTLSLVLLGSSWCGVPATTPALSYNQRIVNGENAVPGS  
WPMQVSLQDNTGFHCGGSLISPNWVYTAHQVTPGRHFHVLGEYDRSSNAEPVQVL  
SIARITHPNWNTMNDLTLLKASPARYTAQVSPVCLATNEALPSGLCTCTGHW  
GRISGVNTPARLQOVPLPLVYNQCRVWGARITDAMICAGSGASSCGGSGGPL  
VCQKNTWLVIGVSWGTKNICNIQAPAMITRVKSFSTWVQVMCYN"  
846..851  
/gene="Ctrl"  
/note="putative"  
868  
/gene="Ctrl"

## gene

## CDS

US-09-735-713a-2 (1-306) x AK009019 (1-868)  
QY 5 MetLeuLeuSerArgAsnLysLeuLeuLeuLeuGlyLeuValPhePheGluArgGly 24  
Db 22 CTACTGCTCAGCCTAACCCCTTAGCCCTGGTCTCCTTGGC-----60  
QY 25 LysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGlnSerLeuValLysVal 44  
Db 61 -----TCTCTCTGGGGCTGGTGTCTCTGCGCAAC 90  
QY 45 GlnProTrpAsnTyrPheAsnLeuPheSerArgLeuGlyGlySerGlnValGluLys 64  
Db 91 AGCCCTGCCTCAGCTAGCTACAAT-----CAGAGATTTGTCAACGGGGAGAAATGCATGTGCCA 144  
QY 65 GlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLys---HisLeuCysGlyGly 83  
Db 145 GGCTCTGGCCCTGCAGGCTGTCTCCAGGATACACCGGCTTCCACTTCTGCGGTGT 204  
QY 84 SerIleValSerProGlnTrpValIleThrAlaAlaHisCys-----IleAlaAsnArg 101  
Db 205 TCTCTCATCAGTCCGCAACTGGGTGTGTCAGCGGTGCCACTGCACAGTCACGCTGGACGC 264  
QY 102 AsnIleValSerLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAspPro 121  
Db 265 CACTTT-----GTCTTTTGGGAGAAATGACCCGATGTCCCAATGCTGAA 309  
QY 122 GlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLysLys 141  
Db 310 CTGTGTCAGGCTCTCTCGATCGCAAGGGCCATCATCTACCTTACTTGGNACGCC---AAC 366  
QY 142 ProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheHisPhe 161  
Db 367 ACCATGAACAATAGCTGACTCTCTCTGAAAGTGTCTGCCAGCCCGGTACACAGCACAA 426  
QY 162 ValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThr 181  
Db 427 GTCTCACCAGTCTGCTGCTTCCACAAACGAGGCACGTGCTTGGGGCTCACCTGTGTGC 486  
QY 182 ThrAlaGlyTrpGlyArgLeuThr---GluGlyGlyValLeuSerGlnValLeuGlnGlu 200  
Db 487 ACCACTGGCTGGGGCCGAATCAGTGGTGGGCAATGTGACACCACTCGCTGGCAGCAA 546  
QY 201 ValAsnLeuProIleLeuThrTrpGluGluCys-----ValAlaLeuLeuThr 217  
Db 547 GTTGTCTTACCCCTGGTCACTGTGAATCAGTGTCCGAGTACTGGGGTGCACGCAATTACC 606  
QY 218 LeuLysArgProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyArg 237  
Db 607 -----GATCCCATGATATATGTGAGGT-----GGCTCAGGCGCC 639  
QY 238 AspAlaCysGlnGlyAspSerGlySerLeuMetCysArgAsnLysLysGlyAlaTrp 257  
Db 640 TCCATCATCGAGGTGACTCAGGAGCCCTCTTCTGCCAG---AAGGAAACACCTGG 696  
QY 258 ThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnVal 277  
Db 697 GTGCTTATTGGGATTTCTCTCTGGGGC-----ACT 726  
QY 278 ArgLysSerAspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrp 297  
Db 727 AAGAACTCAACATACAAGCACCAGCCCATGTACACTCGGGTCAAGCAATTCAGTACCTGG 786

## polyA\_signal

## polyA\_site

```

QY 298 ileHisGlu 300
   |||:||||:
Db 787 ATCAACCAA 795

RESULT 10
BM919200
LOCUS
DEFINITION AGENCOURT_6715624 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5748414
5', mRNA sequence.
ACCESSION BM919200
VERSION BM919200.1 GI:19369579
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1056)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12776 row: c column: 07
High quality sequence stop: 673.
FEATURES
    source
        1..1056
            Location/Qualifiers
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5748414"
                /clone_lib="NIH_MGC_120"
                /lab_host="DH10B"
                /note="Organ: pooled pancreas and spleen; Vector:
                    pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                    source anonymous pool of spleen and pancreas from 28 yo
                    male. Library is oligo-dT primed and directionally cloned
                    (EcoRV site is destroyed upon cloning). Average insert
                    size 1.5 kb, insert size range 1-2.5 kb. Library is
                    normalized and enriched for full-length clones and was
                    constructed by C. Gruber (Invitrogen). Research Genetics
                    tracking code 025. Note: this is a NIH_MGC Library."
BASE COUNT 214 a 323 c 273 g 217 t 29 others
ORIGIN
Alignment Scores:
Pred. No.: 8.88e-36 Length: 1056
Score: 415.50 Matches: 100
Percent Similarity: 51.09% Conservative: 41
Best Local Similarity: 36.23% Mismatches: 96
Query Match: 25.40% Indels: 39
DB: 14 Gaps: 13

US-09-735-713A-2 (1-306)' x BM919200 (1-1056)

QY 36 SerCysGlyGlnSerLeuValGlnProTyrPheAsnIlePheSerArg 55
   |||:||||:
Db 46 TCCTGGGGCTGGCGCATCTCGCATCAACCGGCTAGCTGAGCTTCAGC 99
   |||:||||:

QY 56 IleLeuGlyGlySerGlnValGluLysGlySerTyrProTyrPheGlnValSerLeuLysGln 75
   |||:||||:
Db 100 ATTGTCAACGGGAGATGACAGTGTGGGCTCTGGGCTCGGAGGCTCGCTCGGAGGAC 159
   |||:||||:

QY 76 ArgGlnLys---HisIleCysGlyGlySerIleValSerProGlnTrpValIleThrAla 94
   |||:||||:
Db 160 AGCAGCGGCTTCACATCTTCGGGTGGTGTCTCTCATCAGCCAGCTCCTGGGTGGTCACTGCT 219
   |||:||||:

```

```

QY 95 AlaHisCysIleAlaAsn-----ArgAsnIleValSerThrLeuAsnValThrAlaGly 112
   |||:||||:
Db 220 GCCCACTGCAATGTGAGCCCTGGCGCCATTT-----GTTGCTCTGGGC 264
   |||:||||:

QY 113 GluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrValIle 132
   |||:||||:
Db 265 GAGTATGACCGATCATCAACACGAGAGCCCTTCAGGTTCTGCTCGGCTCTCGGGCCATT 324
   |||:||||:

QY 133 IleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLysMet 152
   |||:||||:
Db 325 ACACACCTAGCTGGAAGTCT---ACCACCATGAACAATGACGTGAGCTGCTGAAGCTC 381
   |||:||||:

QY 153 AlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArgGlu 172
   |||:||||:
Db 382 GCCTGCCAGCCAGTACACAAACAGCATCTCGCCAGTTTGCCTGGCATCTCAACAGGAG 441
   |||:||||:

QY 173 GlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThr---GluGly 191
   |||:||||:
Db 442 GCTCTGACTGAAGCCCTCACGTGTGTCCACCGGCTGGGGTGGCTCAGTGGGGTGGGC 501
   |||:||||:

QY 192 GlyValLeuSerGlnValLeuGlnValAsnLeuProIleLeuThrTrpGluGluCys 211
   |||:||||:
Db 502 AATGTGACACGACCATCTGCAGCAGGTGGCTTGGCCCTGGTCACTGTGAATCACTGC 561
   |||:||||:

QY 212 ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLys-----ThrPhe 227
   |||:||||:
Db 562 -----CGGCAGTACTGGGGCTCAAGTATCACTGACTCCATG 597
   |||:||||:

QY 228 LeuCysThrGlyPheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySer 247
   |||:||||:
Db 598 ATCTGTGAGGT-----GGCGAGGAGTCCCTCGTGCAGGAGTGCCTCGGAGGCGCT 651
   |||:||||:

QY 248 LeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeu 267
   |||:||||:
Db 652 CTTGCTGCCAG---AAGGGAACACATGGGTCTTATTGTGTCTCTGGGGCACC 708
   |||:||||:

QY 268 ---GlyCysGlyArgGlyTrpArgAsnValArgLysSerAspGlnGlySerProGly 286
   |||:||||:
Db 709 AAAAAGTGC-----AATGTGGC-----GCACCTGCT 735
   |||:||||:

QY 287 IlePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIle 302
   |||:||||:
Db 736 GTGTATCTCGAGTTAGCAAGTTTCAGCACCTGGATCAACACAGGTCATA 783
   |||:||||:

RESULT 11
AK014645 1323 bp mRNA linear HTC 19-JAN-2002
LOCUS Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched
DEFINITION library, clone:4733401N09:BRAIN SPECIFIC SERINE PROTEASE-4 TYPE1,
full insert sequence.
ACCESSION AK014645
VERSION AK014645.1 GI:12852632
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (strain:C57BL/6J) 10 days neonate skin cDNA to mRNA,
clone.lib:RIKEN full-length enriched mouse cDNA library
clone:4733401N09.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
AUTHORS Carninci, P. and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
2
REFERENCE
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
JOURNAL prepare full-length cDNA libraries for rapid discovery of new genes
MEDLINE 20499374
PUBMED 11042159

```









CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone Distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: L14M1563 row: p column: 21  
 High quality sequence stop: 815.

FEATURES  
 Location/Qualifiers

1..921  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5224196"  
 /clone\_lib="NIH\_MGC\_120"  
 /lab\_host="DH10B"

/note="Organ: pooled pancreas and spleen; Vector:  
 pCMV-SPOR6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA  
 source anonymous pool of spleen and pancreas from 28 yo  
 male. Library is oligo-dr primed and directionally cloned  
 (EcoRV site is destroyed upon cloning). Average insert  
 size 1.5 kb, insert size range 1-2.5 kb. Library is  
 normalized and enriched for full-length clones and was  
 constructed by C. Gruber (Invitrogen). Research Genetics  
 tracking code 025. Note: this is a NIH\_MGC Library."

BASE COUNT 201 a 278 c 240 g 202 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.57e-35 Length: 921  
 Score: 412.50 Matches: 99  
 Percent Similarity: 50.92% Conservative: 40  
 Best Local Similarity: 36.26% Mismatches: 95  
 Query Match: 25.21% Indels: 39  
 DB: 13 Gaps: 13

US-09-735-713A-2 (1-306) x BI833975 (1-921)

Qy 36 SerCysGlyGlnSerLeuValLysValGlnProTyrPheAsnValPheSerArg 55  
 Db 19 TCCTGGGCGTGGCGATCTCTGCCATCAACCGGCACCTGAGCTTCAGC-----CAGAGG 72  
 Qy 56 IleLeuGlyGlySerGlnValLysGlySerTyrProTyrGlnValSerLeuLysGln 75  
 Db 73 ATGTGCAACGGGGAGATGTCAGTGTGGGCTCTGGCCCTGGCAGGTGCTCCCTCAGGAC 132  
 Qy 76 ArgGlnLys---HisIleCysGlyGlySerIleValSerProGlnTyrValIleThrAla 94  
 Db 133 AGCAGCGCTTCCACTCTCGGTGGTCTCTCATCAGCCAGCTCTGGGTGGTCACTGCT 192  
 Qy 95 AlaHisCysIleAlaAsn-----ArgAsnIleValSerThrLeuAsnValThrAlaGly 112  
 Db 193 GCCCACTGCAATGTGACGCCCTGGCGCCATTTT-----GTTGCTCTGGGC 237  
 Qy 113 GluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrValIle 132  
 Db 238 GAGTATACCGCATCATCAACGACGAGCCCTGTGAGGTCTGTCCGCTCTCTCGGGCCATT 297  
 Qy 133 IleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLysMet 152  
 Db 298 ACACACCTAGCTGGAATCT---ACCACCATGACATGACGTGACGCTCTCAAGCTC 354  
 Qy 153 AlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArgGlu 172  
 Db 355 GCCTGCCGACCGCCAGTACACACGCGATCTCGGCAGTTTCCTGGCATCTCTCAACGAG 414  
 Qy 173 GlnPheGluAlaGlyPheIleCysThrThrAlaGlyTyrGlyArgLeuThr---GluGly 191  
 Db 415 GCTCTGACTGAGGCGCTCAGCTGTGTACACCGCGTGGTGGCTCGCTGGCGGTGGGC 474  
 Qy 192 GlyValLeuSerGlnValLeuGlnValAsnLeuProIleLeuThrTyrPheGluCys 211  
 Db 475 AATGTGACACCGACATCTCAGCAGGTGGCTTTGGCCCTGGTCACTGTCAATCAGTGC 534

Qy 212 ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLys-----ThrPhe 227  
 Db 535 -----CGCAGTACTGGGCTCAAGTATCACTGACTCCCATG 570  
 Qy 228 LeuCysThrGlyPheProAspGlyArgAspAlaCysGlnGlyAspSerGlyGlySer 247  
 Db 571 ATCTGTGCAGGT-----GGCGCAGGTGCTCTCGTCCAGGTGACTCCGGAGGCCCT 624  
 Qy 248 LeuMetCysArgAsnLysLysGlyValAlaTyrThrLeuAlaGlyValThrSerTyrPheLeu 267  
 Db 625 CTCTGTCTGCCAG---AAGGAAACACATGGGTCTTATGTGATGTCTCTGGGGCACC 681  
 Qy 268 ---GlyCysGlyArgGlyTyrPheAsnValArgLysSerAspGlnGlySerProGly 286  
 Db 682 AAAAATGCTC-----AATGTGCGC-----GCACCTGCT 708  
 Qy 287 IlePheThrAspIleSerLysValLeuSerTyrPheHis 299  
 Db 709 GTGTATACTCGATTGAGTTCAGCAAGTTCAGCACCTGGATCAAC 747  
 RESULT 15  
 LOCUS AK007765 876 bp mRNA linear HTC 19-JAN-2002  
 DEFINITION Mus musculus 10 day old male pancreas cDNA, RIKEN full-length  
 enriched library, clone:1810044A17:homolog to CHYMOTRYPSINOGEN B  
 PRECURSOR (EC 3.4.21.1), full insert sequence.  
 ACCESSION AK007765 GI:12841519  
 VERSION AK007765.1  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (strain:C57BL/6J) 10 day old male pancreas cDNA to  
 mRNA, clone\_lib:RIKEN full-length enriched mouse cDNA library  
 clone:1810044A17.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1  
 REFERENCE  
 AUTHORS Carninci,P. and Hayashizaki,Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636  
 2  
 REFERENCE  
 AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,  
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 JOURNAL Genome Res. 10 (10), 1617-1630 (2000)  
 MEDLINE 20499374  
 PUBMED 11042159  
 3  
 REFERENCE  
 AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,  
 Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,  
 Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,  
 Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,  
 Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,  
 Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,  
 Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multipillar sequencer  
 JOURNAL Genome Res. 10 (11), 1757-1771 (2000)  
 MEDLINE 20530913  
 PUBMED 11076861  
 4  
 REFERENCE  
 AUTHORS Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,  
 Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,  
 Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamana,I.,  
 Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,  
 Kadota,K., Matsuda,H., Ashburi,M., Batalov,S., Casavant,T.,  
 Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,  
 Kuehl,P., Lewis,S., Matsuo,Y., Nikaide,I., Pesole,G.,  
 Quackenbush,J., Schriml,L.M., Staibli,F., Suzuki,K., Tomita,M.,  
 Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,  
 Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,



Db 763 GTCACGCCCTCATGCCCCCTGGTTTCAGGAGATCTTGGAA 801

Search completed: March 25, 2003, 02:03:08  
Job time : 1814 secs

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: March 25, 2003, 02:03:14 ; Search time 104 Seconds  
(without alignments)  
2281.689 Million cell updates/sec

Title: US-09-735-713a-2  
Perfect score: 1636  
Sequence: 1 MSLKMLISRNKLLILLGIVF.....IFTDISKVLWIHIQTGN 306

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 538826 seqs, 387737923 residues

Total number of hits satisfying chosen parameters: 1077652

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame\_p2n.model -DEV=xlh  
-Q-/cgn2/uspto.spool/US09735713/runat\_18032003\_124143\_21663/app.query.fasta\_1.455  
-DB-PublishedApplications\_NA -QMT-fastap -SUFFIX-rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS-human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09735713 @CGN 1 1 57 @runat\_18032003\_124143\_21663  
-NCPU=6 -ICPU=3 -NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2.6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2.6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2.6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2.6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2.6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2.6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2.6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2.6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2.6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*
- 10: /cgn2.6/ptodata/1/pubpna/US09\_PUBCOMB.seq.\*
- 11: /cgn2.6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*
- 12: /cgn2.6/ptodata/1/pubpna/US10\_PUBCOMB.seq.\*
- 13: /cgn2.6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*
- 14: /cgn2.6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1626	99.4	921	9	US-09-735-713a-1
2	1626	99.4	1568	9	US-09-735-713a-7
3	1608	98.3	909	9	US-09-735-713a-3
4	1372.5	83.9	1671	10	US-09-888-615-47

5	898	54.9	495	9	US-09-735-713a-5	Sequence 5, Appli
6	490.5	30.0	1327	9	US-09-978-295A-170	Sequence 170, App
7	490.5	30.0	1327	9	US-09-978-697-170	Sequence 170, App
8	490.5	30.0	1327	9	US-09-978-192A-170	Sequence 170, App
9	490.5	30.0	1327	9	US-09-999-832A-170	Sequence 170, App
10	490.5	30.0	1327	9	US-09-978-189-170	Sequence 170, App
11	490.5	30.0	1327	9	US-09-978-608A-170	Sequence 170, App
12	490.5	30.0	1327	9	US-09-978-191A-170	Sequence 170, App
13	490.5	30.0	1327	9	US-09-978-403A-170	Sequence 170, App
14	490.5	30.0	1327	9	US-09-978-564A-170	Sequence 170, App
15	490.5	30.0	1327	9	US-09-978-585A-170	Sequence 170, App
16	490.5	30.0	1327	9	US-10-017-081A-170	Sequence 170, App
17	490.5	30.0	3143	9	US-09-978-295A-168	Sequence 168, App
18	490.5	30.0	3143	9	US-09-978-697-168	Sequence 168, App
19	490.5	30.0	3143	9	US-09-978-192A-168	Sequence 168, App
20	490.5	30.0	3143	9	US-09-999-832A-168	Sequence 168, App
21	490.5	30.0	3143	9	US-09-978-189-168	Sequence 168, App
22	490.5	30.0	3143	9	US-09-978-608A-168	Sequence 168, App
23	490.5	30.0	3143	9	US-09-978-191A-168	Sequence 168, App
24	490.5	30.0	3143	9	US-09-978-403A-168	Sequence 168, App
25	490.5	30.0	3143	9	US-09-978-564A-168	Sequence 168, App
26	490.5	30.0	3143	9	US-09-978-585A-168	Sequence 168, App
27	490.5	30.0	3143	9	US-10-017-081A-168	Sequence 168, App
28	489.5	29.9	2409	10	US-09-888-615-54	Sequence 54, Appl
29	467	28.5	2412	9	US-09-978-295A-68	Sequence 68, Appl
30	467	28.5	2412	9	US-09-978-697-68	Sequence 68, Appl
31	467	28.5	2412	9	US-09-978-192A-68	Sequence 68, Appl
32	467	28.5	2412	9	US-09-999-832A-68	Sequence 68, Appl
33	467	28.5	2412	9	US-09-978-189-68	Sequence 68, Appl
34	467	28.5	2412	9	US-10-174-590-63	Sequence 63, Appl
35	467	28.5	2412	9	US-10-176-758-63	Sequence 63, Appl
36	467	28.5	2412	9	US-10-175-737-63	Sequence 63, Appl
37	467	28.5	2412	9	US-10-173-706-63	Sequence 63, Appl
38	467	28.5	2412	9	US-10-175-738-63	Sequence 63, Appl
39	467	28.5	2412	9	US-10-175-752-63	Sequence 63, Appl
40	467	28.5	2412	9	US-10-176-482-63	Sequence 63, Appl
41	467	28.5	2412	9	US-10-176-757-63	Sequence 63, Appl
42	467	28.5	2412	9	US-10-176-913-63	Sequence 63, Appl
43	467	28.5	2412	9	US-10-180-552-63	Sequence 63, Appl
44	467	28.5	2412	9	US-10-180-557-63	Sequence 63, Appl
45	467	28.5	2412	9	US-10-173-700-63	Sequence 63, Appl

ALIGNMENTS

RESULT 1  
US-09-735-713a-1  
; Sequence 1, Application US/09735713a  
; Patent No. US20020165376A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Turner, C. Alexander Jr.  
; APPLICANT: Abuin, Alejandro  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: No. US20020165376A1e1 Human Proteases and  
; FILE REFERENCE: Polynucleotides Encoding the Same  
; CURRENT APPLICATION NUMBER: US/09/735,713A  
; CURRENT FILING DATE: 2000-12-12  
; PRIOR APPLICATION NUMBER: US 60/171,566  
; PRIOR FILING DATE: 1999-12-22  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 921  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-735-713a-1

Alignment Scores: 1.89e-170 Length: 921  
Pred. No.: 1.89e-170

Score: 1626.00 Matches: 304  
Percent Similarity: 99.35% Conservative: 0  
Best Local Similarity: 99.35% Mismatches: 2  
Query Match: 99.39% Indels: 0  
DB: 9 Gaps: 0

US-09-735-713A-2 (1-306) x US-09-735-713A-1 (1-921)

```
Qy 1 MetSerLeuLysMetLeuLeuSerArgAsnLysLeuLeuLeuLeuGlyIleValPhe 20
Db 1 ATGAGTCTCAAAATGCTTATAGCAGGAAACAGCTGATTTTACTACTAGGAATAGTCTTT 60
Qy 21 PheGluArgGlyLysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGlnSer 40
Db 61 TTTGAACRAGGTAATCTGCARCTCTTTCGGTCCCAAGCTCCAGTGTGGGCAGAGT 120
Qy 41 LeuValLysValGlnProTrpAsnTyrPheAsnIlePheSerArgIleLeuGlyGlySer 60
Db 121 CTGGTTAAGGTACAGCCCTTGGAAATTAATTTAACATTTTCAGTCGCATCTTTGGAGGAAGC 180
Qy 61 GlnValGluLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80
Db 181 CAAGTGGAGAGGGTCTCTATCCCTGGCAGGTATCTCTGAACAAAGCGAGAAGCATATT 240
Qy 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100
Db 241 TGTGAGGAGAGCATGCTCTACCACAGTGGGTGATCATCGGGGCTCAGTCGATTGCAAC 300
Qy 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120
Db 301 AGAAACATTTGTCTACTTTTGAATCTTACTGCTGGAGAGTATGACTTAAGCCAGACACAC 360
Qy 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140
Db 361 CCAGGAGAGCAAACTCTCACATTTGAAACTGTCAATACATACATCCACATTTCTCCACCAAG 420
Qy 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160
Db 421 AAACCAATGGACTATGATATATGCCCCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGCCAC 480
Qy 161 PheValGlyProIleCysLeuProGlnLeuArgGluGlnPheGluAlaGlyPheIleCys 180
Db 481 TTTGTGGGGCCCATATGCTTCCAGAGCTGGGGAGCAATTTGAGGCTGGTATTATTGT 540
Qy 181 ThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValIleuGlnGlu 200
Db 541 ACAACTGAGGCTGGGGCCGCTTAACCTGAAGTGGCGTCTCTCACAAGTCTTGCAGGAA 600
Qy 201 ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg 220
Db 601 GTGAATCTGCCTATTTTTCACCTGGGAGAGTGTGTGGCAGCTCTGTTAACACTAAAGAGG 660
Qy 221 ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys 240
Db 661 CCCATCAGTGGGAAGACCTTTCTTGGCAGAGTTCCTGTGATGGAGGGAGAGACGATGT 720
Qy 241 GlnGlyAspSerGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAla 260
Db 721 CAGGAGATTCAGGAGGTCTACTCATGTCCCGAATAAGAAGGGGCTGGACTTGGCT 780
Qy 261 GlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSer 280
Db 781 GGTGTGACTTCTGGGGTTTGGGCTGTGGTGGAGGCTGGCAACAATGTGAGGANAAGT 840
Qy 281 AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu 300
Db 841 GATCAAGGATCCCTGGGATCTTACAGACATTAAGTAAAGTCTTCTCGATCCACGAA 900
Qy 301 HisIleGlnThrGlyAsn 306
Db 901 CACATCCAACTGGTAAAC 918
```

RESULT 2

US-09-735-713A-7  
; Sequence 7, Application US/09735713A  
; Patent No. US20020165376A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Turner, C. Alexander Jr.  
; APPLICANT: Abuin, Alejandro  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: NO. US20020165376A1el Human Proteases and  
; FILE OF INVENTION: Polynucleotides Encoding the Same  
; FILE REFERENCE: LEX-0108-USA  
; CURRENT APPLICATION NUMBER: US/09/735,713A  
; CURRENT FILING DATE: 2000-12-12  
; PRIOR APPLICATION NUMBER: US 60/171,566  
; PRIOR FILING DATE: 1999-12-22  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 1568  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-735-713A-7

Alignment Scores:  
Pred. No.: 4,14e-170 Length: 1568  
Score: 1626.00 Matches: 304  
Percent Similarity: 99.35% Conservative: 0  
Best Local Similarity: 99.35% Mismatches: 2  
Query Match: 99.39% Indels: 0  
DB: 9 Gaps: 0

US-09-735-713A-2 (1-306) x US-09-735-713A-7 (1-1568)

```
Qy 1 MetSerLeuLysMetLeuLeuSerArgAsnLysLeuLeuLeuLeuGlyIleValPhe 20
Db 245 ATGAGTCTCAAAATGCTTATAGCAGGAAACAGCTGATTTTACTACTAGGAATAGTCTTT 304
Qy 21 PheGluArgGlyLysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGlnSer 40
Db 305 TTTGAACRAGGTAATCTGCARCTCTTTCGGTCCCAAGCTCCAGTGTGGGCAGAGT 364
Qy 41 LeuValLysValGlnProTrpAsnTyrPheAsnIlePheSerArgIleuGlyGlySer 60
Db 365 CTGGTTAAGGTACAGCCCTTGGAAATTAATTTAACATTTTCAGTCGCATCTTGGAGGAAGC 424
Qy 61 GlnValGluLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80
Db 425 CAACTGGAGAAAGGTTCTCTATCCCTGGCAGGTATCTCTGAACAAAGCGAGAAGCATATT 484
Qy 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100
Db 485 TGTGAGGAGAGCATGCTCTACCACAGTGGGTGATTCACGGGGCTCAGTCGATTGCAAC 544
Qy 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120
Db 545 AGAAACATTTGTCTACTTTTGAATGTACTCTGAGAGGTATGACTTAAGCCAGACACAC 604
Qy 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140
Db 605 CCAGGAGAGCAAACTCTCACATTTTGAACACTGTCAATACATCCACATTTCTCCACCAAG 664
Qy 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160
Db 665 AAACCAATGGACTATGATATTTGCCCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGCCAC 724
Qy 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180
Db 725 TTTGTGGGGCCCATATGCTTCCAGAGCTGGGGAGCAATTTGAGGCTGGTATTATTGT 784
Qy 181 ThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnGlu 200
```

```
Db 785 ACAACTGCGAGCTGGGGCGCTTAACCTGAAGTGGCGCTCTCTCACAAAGTCTTGCAGGAA 844
Qy 201 ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg 220
Db 845 GTGAATCTGCCCTATTATGACCTGGAGAGAGTGTGGCAGCTCTGTTAACACTAAAGAGG 904
Qy 221 ProfileSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys 240
Db 905 CCCATCAGTGGGAAGACCTTTCTTTCACAGGTTTCTCGATGAGGAGAGAGCGCATGT 964
Qy 241 GlnGlyaspSerGlyLysLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAla 260
Db 965 CAGGAGATTTCAGAGGTTTCACTATGTGCGGGAATAAGAAAGGGCGCTGGACTCTGGCT 1024
Qy 261 GlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSer 280
Db 1025 GGTGTGACTTCTGGGTTTGGGCTGTGGTCAGGCTGGAGAAACAATGTGAGGAAGAAT 1084
Qy 281 AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu 300
Db 1085 GATCAAGGATCCCTGGGATCTTCACAGACATTAGTAAAGTGCTTCTCTGGATCCACGAA 1144
Qy 301 HisIleGlnThrGlyAsn 306
Db 1145 CACATCCAAACTGGTAAC 1162
RESULT 3
US-09-735-713A-3
; Sequence 3, Application US/09735713A
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abun, Alejandro
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020165376A1el Human Proteases and
; FILE REFERENCE: Polynucleotides Encoding the Same
; CURRENT APPLICATION NUMBER: US/09/735.713A
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,566
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 909
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-713A-3
Alignment Scores:
Pred. No.: 1.8e-168 Length: 909
Score: 1608.00 Matches: 300
Percent Similarity: 99.34% Conservative: 0
Best Local Similarity: 99.34% Mismatches: 2
Query Match: 98.29% Indels: 0
DB: 9 Gaps: 0
US-09-735-713A-2 (1-306) x US-09-735-713A-3 (1-909)
Qy 5 MetLeuIleSerArgAsnLysLeuLeuLeuLeuLeuGlyIleValPhePheGluArgGly 24
Db 1 ATGCTTATACGACGAGCAACAGCTGATTTTACTACTAGGATAGTCTTTTGAACRAGGT 60
Qy 25 LysSerAlaLeuLeuSerLeuProLysAlaProSerCysGlyGlnSerLeuValLysVal 44
Db 61 AAATCTGCARCTCTTTCGCTCCCAAGCTCCAGTGTGGGAGAGTCTGGTTAAGGTA 120
Qy 45 GlnProTrpAsnTrpPheAsnIlePheSerArgIleLeuGlyGlySerGlnValGluLys 64
Db 121 CAGCCTTGGGAATTATTTTAACATTTTCAGTCGCATTTCTTGGAGGAAGCAAGTGGAGAA 180
```

```
Qy 65 GlySerTyPrpTrpGlnValSerLeuLysGlnArgGlnLysHisIleCysGlyGlySer 84
Db 181 GGTTCCTATCCTCGCAGGTATCTCTGAACAAGGCAAGACATATTTGTGGAGGAAGC 240
Qy 85 IleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsnArgAsnIleVal 104
Db 241 ATCGTCTCACACACAGTGGGTGATCAGCGCGCTCACATGCAATGCAAAACAGAAACATTTGTG 300
Qy 105 SerThrLeuAsnValThrAlaGlyGluTyrrAspLeuSerGlnThrAspProGlyGluGln 124
Db 301 TCTACTTTGAATGTACTGTGGAGAGTATGACTTAAGCCAGACAGACCCAGGAGAGCAA 360
Qy 125 ThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLysLysProMetAsp 144
Db 361 ACTCTCACTATTGAACGTGTCATCATCATCCACATTTCTCCACCAAGAAACCAATGAGC 420
Qy 145 TyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHisPheValGlyPro 164
Db 421 TATGATATTGCCCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGCCACTTTTGTGGGGCCC 480
Qy 165 IleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThrAlaGly 184
Db 481 ATATGTCTTCAGAGCTGCGGGAGCAATTTGAGGCTGGTTTTATTTGTACAACTGCAGGC 540
Qy 185 TrpGlyArgLeuThrGluGlyValLeuSerGlnValLeuGlnValValAsnLeuPro 204
Db 541 TGGGGCGCCTTAACCTGAAGTGGCGTCTCTCACAAAGTCTTGCAGGAAGTGAATCTGCCT 600
Qy 205 IleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArgProIleSerGly 224
Db 601 ATTTTGACCTGGGAAGAGTGTGTGCAGCTCTGTTAACACTAAAGAGGCCCATCAGTGGG 660
Qy 225 LysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCysGlnGlyAspSer 244
Db 661 AAGACCTTTCTTTCACAGGTTTTCCTGATGGAGGAGAGACGCGATGTGAGGAGATTCA 720
Qy 245 GlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSer 264
Db 721 GGAGGTTTCATCATGTGCCGAATAAGAAAGGGCGCTGGACTCTGGCTGGTGTGACTTCC 780
Qy 265 TrpGlyLeuGlyCysGlyArgGlyTyrrArgAsnValArgLysSerAspGlnGlySer 284
Db 781 TGGGTTTGGGCTGTGTGGAGGCTGGAGAAACAATGTGAGGAAAGTGTATCAAGGATCC 840
Qy 285 ProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThr 304
Db 841 CCTGGGATCTTTCACAGACATTAGTAAAGTGTCTTCTCTGGATCCAGCAACATCCAAACT 900
Qy 305 GlyAsn 306
Db 901 GGTAAAC 906
RESULT 4
US-09-888-615-47
; Sequence 47, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888.615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 47
```

; LENGTH: 1671  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-888-615-47

## Alignment Scores:

Pred. No.: 4,71e-142 Length: 1671  
Score: 1372.50 Matches: 266  
Percent Similarity: 93.75% Conservative: 4  
Best Local Similarity: 92.36% Mismatches: 6  
Query Match: 83.88% Indels: 12  
DB: 10 Gaps: 2

US-09-735-713a-2 (1-306) x US-09-888-615-47 (1-1671)

```
QY 1 MetSerLeuLysMetLeuLysSerArgAsnLysLeuLysLeuLeuLeuGlyLeuValPhe 20
Db 1 ATGAGTCTCAAAATGCTTATAGCAGGAACAAGCTGATTTACTACTAGGAATAGTCTTTT 60
QY 21 PheGluArgGlyLysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGlnSer 40
Db 61 TTTGAACGAGGTAATCTCAACTCTTCGCTCCCAAGCTCCCAAGTTGTGGCAGAGT 120
QY 41 LeuValLysValGlnProTrpAsnTrpPheAsnIlePheSerArgIleLeuGlyGlySer 60
Db 121 CTGGTTAAGGTACAGCTTTGGAATATATTTAAACATTTTCAGTCGCATTTCTGGAGGAAGC 180
QY 61 GlnValGluLysGlySerTrpProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80
Db 181 CAAGTGGAGAGGGTTCCTATCCCTGGCAGGATCTCTGAAACAAAGGAGGAGCATATT 240
QY 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100
Db 241 TGTGGAGGAAGCATCTCTCACACAGTGGGTGATCACGCGGCTCAGTCATTCGAAC 300
QY 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTrpAspLeuSerGlnThrAsp 120
Db 301 AGAAACATTGTGTCTACTTTGAATGTACTGCTGGAGAGATGACTTAAGCCAGACAGAC 360
QY 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140
Db 361 CCAGGAGAGCAAACTCTACTATTGAACTGTCATCATACATCCACATTTCTCCACCAAG 420
QY 141 LysProMetAspTrpAspIleAlaLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160
Db 421 AAACCAATGGACTATGATATGCGCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGCCAC 480
QY 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180
Db 481 TTTGTGGGCGCCATATGCTTCCAGAGCTGCGGGAGCAATTTGAGGCTGGTTTATTTGT 540
QY 181 ThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnGlu 200
Db 541 ACAACTGACAGGCTGGGCGCGCTTAAGTGAAGTGGCGCTCTCTCACAGTCTTGAGGAA 600
QY 201 ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg 220
Db 601 GTGAATCTGCCATTTTGAACCTGGGAAGAGTGTGTGGCAGCTCTGTTAACACTAAAGAGG 660
QY 221 ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys 240
Db 661 CCCATCAGTGGGAAGACCTTTCTTGCACAGGTTTTCCTGATGGAGGGAGACGCATGT 720
QY 241 GlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrp-ThrLeuAl 260
Db 721 CAGGAGAGATTACGAGGTTCACTATGTGCCGGAATAAGAAAGGGCCCTGGGACTCTGCC 780
QY 260 aglyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSe 280
Db 781 TGGT-----CAATTGGG-----AGGCTCAGGTGGGA 807
QY 280 rAspGlnGlySerProGlyIle 287
Db 287 TTTT-----TTTTTTTTTTT-----
```

Db 808 GGATCGCTTTGAGTCCAGGAGTT 829

## RESULT 5

US-09-735-713a-5

```
; Sequence 5, Application US/09735713A
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abuin, Alejandro
; APPLICANT: Friedrich, Glenn
; APPLICANT: zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0108-USA
; CURRENT APPLICATION NUMBER: US/09/735,713A
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,566
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-735-713a-5
```

## Alignment Scores:

Pred. No.: 2.13e-90 Length: 495  
Score: 898.00 Matches: 164  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 54.89% Indels: 0  
DB: 9 Gaps: 0

US-09-735-713a-2 (1-306) x US-09-735-713a-5 (1-495)

```
QY 143 MetAspTrpAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHisPheVal 162
Db 1 ATGGACTATGATATATGCGCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCACTTTGTG 60
QY 163 GlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThr 182
Db 61 GGGCCCATATGCTCTCCAGAGCTGCGGAGCAATTTGAGGCTGGTTTTATTGTACAAC 120
QY 183 AlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnValAsn 202
Db 121 GCAGGCTGGGGCGGCTTAAGTGGCGTCTCTCAAGTCTTCAGGAAGTGAAT 180
QY 203 LeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArgProIle 222
Db 181 CTGCTATTTTTCACCTGGGAAGAGTGTGTGGCAGCTCTCTTAACACTAAGAGGCCCATC 240
QY 223 SerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCysGlnGly 242
Db 241 AGTGGGAAGACCTTTCTTTCACAGGTTTCTGTATGGAGGGAGAGACGCATGTGAGGGA 300
QY 243 AspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAlaGlyVal 262
Db 301 GATTTCAGGAGGTTCACTCATGTGCCGGAATAAGAAAGGGGCGCTGGACTCTGGCTGGTGTG 360
QY 263 ThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSerAspGln 282
Db 361 ACTTCTGGGGTTTGGGCTGTGTGGCTGGAGGCTGGAGAACAAATGTGAGGAAAAGTGAACA 420
QY 283 GlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIle 302
Db 421 GGATCCCTGGGATCTTTCACAGACATTAGTAAAGTGTCTTCCTGGATCCACGAACATC 480
QY 303 GlnThrGlyAsn 306
Db 303 TTTTTTTTTT
QY 481 CAAACTGGGTAAC 492
```

RESULT 6  
US-09-978-295A-170  
; Sequence 170, Application US/09978295A  
; Patent No. US20020156006A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan  
; APPLICANT: Ferrara, Napoleon  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Kuo, Sophia S.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James;  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Shelton, David L.  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2630PIC11  
; CURRENT APPLICATION NUMBER: US/09/978,295A  
; CURRENT FILING DATE: 2001-10-15  
; PRIOR APPLICATION NUMBER: 09/918585  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/064249  
; PRIOR FILING DATE: 1997-11-03  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066364  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: 60/077450  
; PRIOR FILING DATE: 1998-03-10  
; PRIOR APPLICATION NUMBER: 60/077632  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077641  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077649  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077791  
; PRIOR FILING DATE: 1998-03-12  
; PRIOR APPLICATION NUMBER: 60/078004  
; PRIOR FILING DATE: 1998-03-13  
; PRIOR APPLICATION NUMBER: 60/078886  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078936  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078939  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079294  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: 60/079656  
; PRIOR FILING DATE: 1998-03-26  
; PRIOR APPLICATION NUMBER: 60/079664  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079689  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079663  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079728  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079786  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079920  
; PRIOR FILING DATE: 1998-03-30  
; PRIOR APPLICATION NUMBER: 60/079923  
; PRIOR FILING DATE: 1998-03-30  
; PRIOR APPLICATION NUMBER: 60/080105  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080107  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080165  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080194  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080327  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080328  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080333  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080334  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/081070  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081049  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081071  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081195  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081203  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081229  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081955  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081817  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081819  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081952  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081838  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/082568  
; PRIOR FILING DATE: 1998-04-21  
; PRIOR APPLICATION NUMBER: 60/082569  
; PRIOR FILING DATE: 1998-04-21  
; PRIOR APPLICATION NUMBER: 60/082704  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082804  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082700  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082797  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082796  
; PRIOR FILING DATE: 1998-04-23  
; PRIOR APPLICATION NUMBER: 60/083336  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/083392  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083495  
; PRIOR FILING DATE: 1998-04-29

;; PRIOR APPLICATION NUMBER: 60/083496  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083499  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083545  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083554  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083558  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083559  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083500  
;; PRIOR FILING DATE: 1998-04-29  
;; PRIOR APPLICATION NUMBER: 60/083742  
;; PRIOR FILING DATE: 1998-04-30  
;; PRIOR APPLICATION NUMBER: 60/084366  
;; PRIOR FILING DATE: 1998-05-05  
;; PRIOR APPLICATION NUMBER: 60/084414  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084441  
;; PRIOR FILING DATE: 1998-05-06  
;; PRIOR APPLICATION NUMBER: 60/084637  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084639  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084640  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084598  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084600  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084643  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085582  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085700  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085689  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085580  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085573  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:  
Pred. No.: 9, 92e-45  
Score: 490.50  
Percent Similarity: 56.42%  
Best Local Similarity: 38.91%  
Query Match: 29.98%  
DB: 9

US-09-735-713a-2 (1-306) x US-09-978-295A-170 (1-1327)

Qy 54 SerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu 73  
Db 534 AGCCCATGTTGGTGGAGTGTGCTCCCGAGGGTGTGTCCTCCGAGGCGGCGAGCCAGCCCTC 593  
Qy 74 LysGlnArgGlnLysHisIleCysGlyGlySerLeuValSerProGlnTrpValIleThr 93

Db 594 CAGGTTGGGGTCCGACACATCTGTGGGGGGCCCTCATCGCTGACCGCTGGGTGATAACA 653  
Qy 94 AlaAlaHisCysIleAlaAsnArgAsnIleValSerThrLeu-----AsnValThrAla 111  
Db 654 GCTGCCCACTGCTTCCAGGAGGACAGCATGTGCTCCACGGTGTGTTGGACCGTGTTCCTCG 713  
Qy 112 GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131  
Db 714 GCAAGGTGTGCACAACTCGCGCTGGCGCTGGAGAGGTGCTCTCAAGGTGAGCCCGCTCG 773  
Qy 132 IleIleHisProHisPheSerThrLysLysPrometAspTyrAspIleAlaLeuLeuLys 151  
Db 774 CTCCTGCACCCG---TACCACGAAGAGGACACCCATGACTACGAGTGGCGCTGCTGCAG 830  
Qy 152 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171  
Db 831 CTCGACACCCCGTGTGCTCGCGCCCGCTGCGCCCGCTGCTGCTGCCCGCGCTCC 890  
Qy 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTyrGlyArgLeuThrGluGly 191  
Db 891 CACTTCTTCGAGCCCGGCTGCACTGCTGGATTACGGCTGGGGCGCTTGGCGAGGCG 950  
Qy 192 GlyValLeuSerGlnValLeuGlnValAlaGlyValLeuProIleLeuThrTrpGluLucys 211  
Db 951 GGCCCATCAGCAGCGCTCTGCAGAAAGTGGATGTGCAGTTGATCCACAGGACCTGTGC 1010  
Qy 212 ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly 231  
Db 1011 AGCGAGGCTATTCGCTACCGGTGACGCCA-----CGCATGCTGTGTGCCGCG 1058  
Qy 232 PheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg 251  
Db 1059 TACCCCAAGGGCAAGAGGATGCTGTGAGGTGACTCAGGTGGTCCGCTGGTGTGCAAG 1118  
Qy 252 AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg 271  
Db 1119 GCACCTCAGTGGCGCTGCTTCTGCGGGGCTGGTGCAGCTGGGGCTGTGGCTGTGCCGG 1178  
Qy 272 GlyTrpArgAsnAsnValArgLysSerAspGlnGlySerPro-----GlyIlePhe 288  
Db 1179 -----CCTAACTACTTGGCGCTGTAC 1199

RESULT 7  
US-09-978-697-170  
; Sequence 170, Application US/09978697  
; Patent No. US20020169284A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan  
; APPLICANT: Ferrara, Napoleon  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Kuo, Sophia S.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James;  
; APPLICANT: Paoni, Nicholas F.

APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630P1C27  
CURRENT APPLICATION NUMBER: US/09/978,697  
CURRENT FILING DATE: 2001-10-16  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084639  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640

```

; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:
Pred. No.:          9,928-45      Length:      1327
Score:             490.50        Matches:    100
Percent Similarity: 56.42%       Conservative: 45
Best Local Similarity: 38.91%    Mismatches: 89
Query Match:       29.98%        Indels:     23
DB:                9            Gaps:         5

US-09-735-713A-2 (1-306) x US-09-978-697-170 (1-1327)

QY 54 SerAtgIleuGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu 73
Db 534 ACCCGATTTGGTGGAGCATCTGTCCCGAGGGTGTAGTGGCCATGGCAGCCACGCTC 593

QY 74 LysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpValIleThr 93
Db 594 CAGGTTCGGGGTCGACACATCTGTGGGGGGCCCTCATCGCTGACCGTGGGTGATAACA 653

QY 94 AlaAlaHisCysIleAlaAsnArgAsnIleValSerThrLeu-----AsnValThrAla 111
Db 654 GCTGCCACTGCTCCAGGAGGACAGCATGGCTCCACGGTGTCTGGACCGTGTTCCTG 713

QY 112 GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131
Db 714 GCAAGGTGTGCACAACTCGCGCTGGCTGGAGAGGTGTCTTCAAGGTGAGCGCGCTG 773

QY 132 IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLys 151
Db 774 CTCCTGCACCCG---TACCACGAGAGGACACCCATGACTACGAGTGGCGCTGTGCGAG 830

QY 152 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171
Db 831 CTCGACCCCGGTGTGGCTCGCGCGCGCTGGCCCGCTGTGCTGCCCGCGCGCTCC 890

QY 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly 191
Db 891 CACTTCTTGAGCCCGCGCTGCACACTGCTGGATFACGGCTGGGGCGCTTGGCGAGGCG 950

QY 192 GlyValLeuSerGlnValLeuGlnValAsnLeuProIleLeuThrTrpGluGluCys 211
Db 951 GCGCCCATCAGCAGCTCTCGAAGAAGTGTGATGTGCAGTTGATCCCAAGGACCTGTGC 1010

QY 212 ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly 231
Db 1011 AGCGAGGCTATCGTACCAGGTGACGCCA-----CGCATGCTGTGTGCCGCG 1058

QY 232 PheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg 251
Db 1059 TACCGCAAGGCAAGAGGATCGCTCAGGGTGACTCAGGTGGTCCGCTGGTGTGCAAG 1118

QY 252 AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg 271
Db 1119 GCACCTAGTGGCGCTGGTTCCTGGCGGGCTGGTCAAGTGGGCTGGCTGGTGGCCGG 1178

QY 272 GlyTrpArgAsnAsnValArgLysSerAspGlnGlySerPro-----GlyIlePhe 288
Db 1179 -----CCTAACTACTTCGGCGCTTAC 1199

QY 289 ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly 305
Db 1200 ACCCGCATCACAGGTGTGATCATGCTGGATCCAGCAAGTGGTGGTACCTGAGGA 1250

RESULT 8
US-09-978-192A-170
; Sequence 170, Application US/09978192A
; Patent No. US2002017755A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C9
; CURRENT APPLICATION NUMBER: US/09/978,192A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
```

>	PRIOR FILING DATE:	1998-04-21
>	PRIOR APPLICATION NUMBER:	60/082704
>	PRIOR FILING DATE:	1998-04-22
>	PRIOR APPLICATION NUMBER:	60/082804
>	PRIOR FILING DATE:	1998-04-22
>	PRIOR APPLICATION NUMBER:	60/082700
>	PRIOR FILING DATE:	1998-04-22
>	PRIOR APPLICATION NUMBER:	60/082797
>	PRIOR FILING DATE:	1998-04-22
>	PRIOR APPLICATION NUMBER:	60/082796
>	PRIOR FILING DATE:	1998-04-23
>	PRIOR APPLICATION NUMBER:	60/083336
>	PRIOR FILING DATE:	1998-04-27
>	PRIOR APPLICATION NUMBER:	60/083322
>	PRIOR FILING DATE:	1998-04-28
>	PRIOR APPLICATION NUMBER:	60/083392
>	PRIOR FILING DATE:	1998-04-29
>	PRIOR APPLICATION NUMBER:	60/083495
>	PRIOR FILING DATE:	1998-04-29
>	PRIOR APPLICATION NUMBER:	60/083496
>	PRIOR FILING DATE:	1998-04-29
>	PRIOR APPLICATION NUMBER:	60/083499
>	PRIOR FILING DATE:	1998-04-29
>	PRIOR APPLICATION NUMBER:	60/083545
>	PRIOR FILING DATE:	1998-04-29
>	PRIOR APPLICATION NUMBER:	60/083554
>	PRIOR FILING DATE:	1998-04-29
>	PRIOR APPLICATION NUMBER:	60/083558
>	PRIOR FILING DATE:	1998-04-29
>	PRIOR APPLICATION NUMBER:	60/083559
>	PRIOR FILING DATE:	1998-04-29
>	PRIOR APPLICATION NUMBER:	60/083500
>	PRIOR FILING DATE:	1998-04-29
>	PRIOR APPLICATION NUMBER:	60/083742
>	PRIOR FILING DATE:	1998-04-30
>	PRIOR APPLICATION NUMBER:	60/084366
>	PRIOR FILING DATE:	1998-05-05
>	PRIOR APPLICATION NUMBER:	60/084414
>	PRIOR FILING DATE:	1998-05-06
>	PRIOR APPLICATION NUMBER:	60/084441
>	PRIOR FILING DATE:	1998-05-06
>	PRIOR APPLICATION NUMBER:	60/084637
>	PRIOR FILING DATE:	1998-05-07
>	PRIOR APPLICATION NUMBER:	60/084639
>	PRIOR FILING DATE:	1998-05-07
>	PRIOR APPLICATION NUMBER:	60/084640
>	PRIOR FILING DATE:	1998-05-07
>	PRIOR APPLICATION NUMBER:	60/084598
>	PRIOR FILING DATE:	1998-05-07
>	PRIOR APPLICATION NUMBER:	60/084600
>	PRIOR FILING DATE:	1998-05-07
>	PRIOR APPLICATION NUMBER:	60/084627
>	PRIOR FILING DATE:	1998-05-07
>	PRIOR APPLICATION NUMBER:	60/084643
>	PRIOR FILING DATE:	1998-05-07
>	PRIOR APPLICATION NUMBER:	60/085339
>	PRIOR FILING DATE:	1998-05-13
>	PRIOR APPLICATION NUMBER:	60/085338
>	PRIOR FILING DATE:	1998-05-13
>	PRIOR APPLICATION NUMBER:	60/085323
>	PRIOR FILING DATE:	1998-05-13
>	PRIOR APPLICATION NUMBER:	60/085582
>	PRIOR FILING DATE:	1998-05-15
>	PRIOR APPLICATION NUMBER:	60/085579
>	PRIOR FILING DATE:	1998-05-15
>	PRIOR APPLICATION NUMBER:	60/085580
>	PRIOR FILING DATE:	1998-05-15
>	PRIOR APPLICATION NUMBER:	60/085573
>	PRIOR FILING DATE:	1998-05-15

; PRIOR APPLICATION NUMBER: 60/085704  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085697

## Alignment Scores:

Pred. No.: 9,92e-45 Length: 1327  
Score: 490.50 Matches: 100  
Percent Similarity: 56.42% Conservative: 45  
Best Local Similarity: 38.91% Mismatches: 89  
Query Match: 29.98% Indels: 23  
DB: 9 Gaps: 5

US-09-735-713A-2 (1-306) x US-09-978-192A-170 (1-1327)

Qy 54 SerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu 73  
Db 534 AGCCGCAATGTTGGAGCTGTCTCCGAGGGTGTAGTGGCCATGCGAGCCAGCCCTC 593  
Qy 74 LysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpValIleThr 93  
Db 594 CAGGTTCGGGTGCGACACATCTGTGGGGGGCCCTCATCGCTGACCGCTGGGTGATAACA 653  
Qy 94 AlaAlaHisCysIleAlaAsnArgAsnIleValSerThrLeu-----AsnValThrAla 111  
Db 654 GCTGCCCACTGCTCCAGAGGAGACAGCATGGCCCTCCAGCGGTGTGTGACCGTGTCTG 713  
Qy 112 GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131  
Db 714 GCGAGGTGTGGCAACCTCGCGTGGCTGGAGAGGTGTCTTCAAGTGTGAGCGCGCTG 773  
Qy 132 IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLys 151  
Db 774 CTCCTGCACCGG---TACCAGGAAGAGGACAGCCATGACTACGAGCTGGCGCTGTGCAG 830  
Qy 152 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171  
Db 831 CTCGACCCCGGTGTGCGCTGCGCCCGCTGCGCCCGCTGCTGCGCGCGCTGCC 890  
Qy 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTyrGlyArgLeuThrGluGly 191  
Db 891 CACTTCTCGAGCCGGCTGCCTGTGGATTACGGGCTGGGGCGCTTCCGGGAGGCGC 950  
Qy 192 GlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTrpGluGluCys 211  
Db 951 GGCCCATCAGCAACGCTCTCAGAAAGTGTGTGCAAGTGTGATCCACAGGACTGTGC 1010  
Qy 212 ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly 231  
Db 1011 AGCGAGGCTTATCGCTACCGGTGACGCCA-----CGCATGCTGTGTGCCGCGC 1058  
Qy 232 PheProAspGlyGlyArgAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg 251  
Db 1059 TACCCGACGAGCAAGAGTGCCTGTGAGGTGACTCAGTGTGCTGCTGCTGTGCAG 1118  
Qy 252 AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg 271  
Db 1119 GCACCTAGTGGCGCTGTCTGCGGGGCTGTGCTGAGTGGCGCTGGGTGTGGCGG 1178  
Qy 272 GlyTrpArgAsnValArgLysSerAspGlnGlySerPro-----GlyIlePhe 288  
Db 1179 -----CCTAACTACTTCGGCGCTTAC 1199  
Qy 289 ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly 305  
Db 1200 ACCCGCATCAGGTGTGTGATCAGTGGATCCAGCAAGTGTGTACCTGAGGA 1250

## RESULT 9

US-09-999-832A-170  
; Sequence 170, Application US/09999832A  
; Publication No. US20020192706A1  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker Kevin P.

; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan  
; APPLICANT: Ferrara, Napoleon  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Kuo, Sophia S.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Shelton, David L.  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE REFERENCE: P2630PLC63  
; CURRENT APPLICATION NUMBER: US/09/999,832A  
; CURRENT FILING DATE: 2001-10-24  
; PRIOR APPLICATION NUMBER: 09/918585  
; PRIOR FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/064249  
; PRIOR FILING DATE: 1997-11-03  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066364  
; PRIOR FILING DATE: 1997-11-21  
; PRIOR APPLICATION NUMBER: 60/077450  
; PRIOR FILING DATE: 1998-03-10  
; PRIOR APPLICATION NUMBER: 60/077632  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077641  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077649  
; PRIOR FILING DATE: 1998-03-11  
; PRIOR APPLICATION NUMBER: 60/077791  
; PRIOR FILING DATE: 1998-03-12  
; PRIOR APPLICATION NUMBER: 60/078004  
; PRIOR FILING DATE: 1998-03-13  
; PRIOR APPLICATION NUMBER: 60/078886  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078936  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/078939  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/079294  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: 60/079656  
; PRIOR FILING DATE: 1998-03-26  
; PRIOR APPLICATION NUMBER: 60/079664  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079689  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079663  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079728  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079786

; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: 60/079920  
; PRIOR FILING DATE: 1998-03-30  
; PRIOR APPLICATION NUMBER: 60/079923  
; PRIOR FILING DATE: 1998-03-30  
; PRIOR APPLICATION NUMBER: 60/080105  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080107  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080165  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080194  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 60/080327  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080328  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080333  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/080334  
; PRIOR FILING DATE: 1998-04-01  
; PRIOR APPLICATION NUMBER: 60/081070  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081049  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081071  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081195  
; PRIOR FILING DATE: 1998-04-08  
; PRIOR APPLICATION NUMBER: 60/081203  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081229  
; PRIOR FILING DATE: 1998-04-09  
; PRIOR APPLICATION NUMBER: 60/081955  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081817  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081819  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081952  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/081838  
; PRIOR FILING DATE: 1998-04-15  
; PRIOR APPLICATION NUMBER: 60/082568  
; PRIOR FILING DATE: 1998-04-21  
; PRIOR APPLICATION NUMBER: 60/082569  
; PRIOR FILING DATE: 1998-04-21  
; PRIOR APPLICATION NUMBER: 60/082704  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082804  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082700  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082797  
; PRIOR FILING DATE: 1998-04-22  
; PRIOR APPLICATION NUMBER: 60/082796  
; PRIOR FILING DATE: 1998-04-23  
; PRIOR APPLICATION NUMBER: 60/083336  
; PRIOR FILING DATE: 1998-04-27  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/083392  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083495  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083496  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083499  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083545  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083554  
; PRIOR FILING DATE: 1998-04-29

; PRIOR APPLICATION NUMBER: 60/083558  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083559  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083500  
; PRIOR FILING DATE: 1998-04-29  
; PRIOR APPLICATION NUMBER: 60/083742  
; PRIOR FILING DATE: 1998-04-30  
; PRIOR APPLICATION NUMBER: 60/084366  
; PRIOR FILING DATE: 1998-05-05  
; PRIOR APPLICATION NUMBER: 60/084414  
; PRIOR FILING DATE: 1998-05-06  
; PRIOR APPLICATION NUMBER: 60/084441  
; PRIOR FILING DATE: 1998-05-06  
; PRIOR APPLICATION NUMBER: 60/084637  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084639  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084640  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084598  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084627  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/084643  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/085339  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/085338  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/085323  
; PRIOR FILING DATE: 1998-05-13  
; PRIOR APPLICATION NUMBER: 60/085582  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085700  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085689  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085579  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085580  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085573  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085704  
; PRIOR FILING DATE: 1998-05-15  
; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:

Pred. No.: 9.92e-45 Length: 1327  
Score: 490.50 Matches: 100  
Percent Similarity: 56.42% Conservative: 45  
Best Local Similarity: 38.91% Mismatches: 89  
Query Match: 29.98% Indels: 23  
DB: 5 Gaps: 5

US-09-735-713a-2 (1-306) x US-09-999-832a-170 (1-1327)

QY 54 SerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu 73  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 534 AGCCGCATTGTGGTGAGCTGTCTCCGAGGGTGAGTGGCCATGGCAGCCGAGCCTC 593  
QY 74 LysGlnArgGlnLysHisIleCysGlySerIleValSerProGlnTrpValIleThr 93  
::: ||| ::::::::::::::::::::: ::::::::::::::::::::: :::::::::::::::::::::  
Db 594 CAGGTTCCGGGTCCGACACATCTGTGGGGGGCCCTCATCGCTGACCGCTGGGTGATACA 653  
QY 94 AlaAlaHisCysIleAlaAsnArgAsnIleValSerThrLeu-----AsnValThrAla 111  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 654 GCTGCCCACTGCTTCCAGGAGGACAGCATGGCCCTCCACGGTGTGTGACCGCTGTCTCTG 713  
QY 112 GlyCluTyrAspLeuSerGlnThrAspProGlyCluGlnThrLeuThrIleGluThrVal 131

```
Db 714 GGCAAGGTGGCAGACTCGCCTGGCTGGAGAGGTTCCTTAAGGTGAGCCGCCGTG 773
; FILE REFERENCE: P2630PIC7
; CURRENT APPLICATION NUMBER: US/09/978,189
QY 132 IleIleHisProHisSerThrLysLysPrometAspTyrAspIleAlaLeuLeuLys 151
; PRIOR FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
Db 774 CTCCTGCACCG--TACCACGAAGAGCAGACCATGACTACGACGTGGCGTGTGCAG 830
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
QY 152 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
Db 831 CTCGACCACCCGGTGTGGCTGGCGCGCGCTGCTGCGCGCGCGCTCC 890
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
QY 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpClyArgIleThrGluGly 191
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/076364
Db 891 CACTTCTTGAGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 950
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077450
QY 192 GlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTrpGluGluCys 211
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077632
Db 951 GGCCCATCAGCAACCTCTGCAGAAAGTGGATGTCAGTTCATCCACAGACCTGTGC 1010
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
QY 212 ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly 231
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
Db 1011 AGCGAGGCTATCGCTACCAGGTGAGCGCA-----CGCATGCTGTGTGCGCGC 1058
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/077791
QY 232 PheProAspGlyClyArgAspAlaCysGlnGlyAspSerGlyClySerLeuMetCysArg 251
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078004
Db 1059 TACCGCAAGGCAAGAAAGATGCTGTAGGGTACTCAGGTGGTCCGCTGGTGTGCAAG 1118
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078886
QY 252 AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg 271
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
Db 1119 GCACCTCAGTGGCGCTGTCTCTGGCGGGCTGGTCACTGAGTGGCGCTGGTGGCGG 1178
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
QY 272 GlyTrpArgAsnValArgLysSerAspGlnGlySerPro-----GlyIlePhe 288
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
Db 1179 -----CTAACTACTTCGGCGCTACT 1199
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
QY 289 ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly 305
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
Db 1200 ACCCGCATCAGGTGTGTATCAGTGGATCCAGCAAGTGGTACCTGAGGA 1250
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; TITLE OF INVENTION: Acids Encoding the Same
```

## RESULT 10

US-09-978-189-170

; Sequence 170, Application US/09978189

; Publication No. US20030004102A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Baker Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan

; APPLICANT: Ferrara, Napoleon

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Kuo, Sophia S.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Shelton, David L.

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same



594	Db	CAGGTCGGGGTGCAGACATCTGTGGGGGGGCCCTCATCGCTGACCGCTGGGTGATACA	655
94	Qy	AlaAlaHisCysIleAlaAsnArgAsnIleValSerThrLeu-----AsnValThrAla	111
654	Db	CTGCGCCACTGCTCCAGGAGGACATGGCTCCAGGTGCTGTGGACCGTGTTCGT	713
112	Qy	GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal	131
714	Db	GGCAAGGTGTGGCAGAACTCGCGCTGGCCTGGAGAGGTGCTCTCAAGGTGAGCGCGCTG	773
132	Qy	IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLys	151
774	Db	CTCTGTCACCG---TACCACGAAGAGACAGCACTAGCAGCTGGCGCTGCTCGAG	830
152	Qy	MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg	171
831	Db	CTGACACCCGGTGTGGCTGGCGCGCTGGCCCGCTGCTGCTGCCCGCGCGCTCC	890
172	Qy	GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTyrGlyArgLeuThrGluGly	191
891	Db	CACCTTCTCGAGCCCGCTGCCTGCTGGATTACGGGCTGGGGCGCTTGCAGGAGGC	950
192	Qy	GlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTyrGluGluCys	211
951	Db	GGCCCATCAGCAACCGCTCTGCAGAAAGTGGATGTGCAGTTGATCCACAGAGCTGTGC	1010
212	Qy	ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly	231
1011	Db	AGCGAGCGCTATCGCTACCAGGTGAGCCA-----CGCATGCTGTGTGCCGCG	1058
232	Qy	PheProAspGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg	251
1059	Db	TACCGCAAGGCGAAGAAGATCGCTGTCAAGGTGACTCAGGTGGTCCGCTGGTGTGCAAG	1118
252	Qy	AsnLysLysGlyAlaTyrThrLeuAlaGlyValThrSerTyrGlyLeuGlyCysGlyArg	271
1119	Db	GCACCTCAGTGGCGCTGGTTCCTCGGGGCTGGTCAGCTGGGGCTGGGCTGGCGCG	1178
272	Qy	GlyTyrArgAsnAsnValArgLysSerAspGlnGlySerPro-----GlyIlePhe	288
1179	Db	-----CCTAATACTTCCGGCTGTAC	1199
289	Qy	ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly	305
1200	Db	ACCGCATCAGAGTGTGATCATCGTGGATCCAGCAAGTGTGACCTGAGGA	1250
RESULT 12			
US-09-978-191A-170			
; Sequence 170, Application US/09978191A			
; Publication No. US20030050239A1			
; GENERAL INFORMATION:			
; APPLICANT: Ashkenazi, Avi			
; APPLICANT: Baker Kevin P.			
; APPLICANT: Botstein, David			
; APPLICANT: Desnoyers, Luc			
; APPLICANT: Eaton, Dan			
; APPLICANT: Ferrara, Napoleon			
; APPLICANT: Filvaroff, Ellen			
; APPLICANT: Fong, Sherman			
; APPLICANT: Gao, Wei-Qiang			
; APPLICANT: Gerber, Hanspeter			
; APPLICANT: Gerritsen, Mary E.			
; APPLICANT: Goddard, Audrey			
; APPLICANT: Godowski, Paul J.			
; APPLICANT: Grimaldi, J. Christopher			
; APPLICANT: Gurney, Austin L.			
; APPLICANT: Hillan, Kenneth J			
; APPLICANT: Kljavin, Ivar J.			
; APPLICANT: Kuo, Sophia S.			
; APPLICANT: Napier, Mary A.			
; APPLICANT: Pan, James:			
; APPLICANT: Paoni, Nicholas F.			
; APPLICANT: Roy, Margaret Ann			

APPLICANT: Shelton, David L.	PRIOR FILING DATE: 1998-04-01
APPLICANT: Stewart, Timothy A.	PRIOR APPLICATION NUMBER: 60/081070
APPLICANT: Tumas, Daniel	PRIOR FILING DATE: 1998-04-08
APPLICANT: Williams, P. Mickey	PRIOR APPLICATION NUMBER: 60/081049
APPLICANT: Wood, William I.	PRIOR FILING DATE: 1998-04-08
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic	PRIOR APPLICATION NUMBER: 60/081071
FILE REFERENCE: P2630PIC4	PRIOR FILING DATE: 1998-04-08
CURRENT APPLICATION NUMBER: US/09/978,191A	PRIOR APPLICATION NUMBER: 60/081195
CURRENT FILING DATE: 2001-10-15	PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 09/918585	PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 2001-07-30	PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/062250	PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1997-10-17	PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/064249	PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1997-11-03	PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/065311	PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1997-11-13	PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/066364	PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1997-11-21	PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/077450	PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-03-10	PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/077632	PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-03-11	PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/077791	PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-03-12	PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/078004	PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-03-13	PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/078886	PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-03-20	PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/078936	PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-03-20	PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/078910	PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-03-20	PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/078939	PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-03-20	PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/079294	PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-03-25	PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/079656	PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-03-26	PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/079664	PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-03-27	PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/079689	PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-03-27	PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/079663	PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-03-27	PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/079728	PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-03-27	PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/079786	PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-03-27	PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/079920	PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-03-30	PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/079923	PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-03-30	PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/080105	PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-03-31	PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/080107	PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-03-31	PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/080165	PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-03-31	PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/080194	PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-03-31	PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/080327	PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-04-01	PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/080328	PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-04-01	PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/080333	PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-04-01	PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/080334	PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-04-01	PRIOR FILING DATE: 1998-05-07

```
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:
Pred. No.:          9,92e-45          Length:      1327
Score:              490.50            Matches:    100
Percent Similarity: 56.42%            Conservative: 45
Best Local Similarity: 38.91%          Mismatches:  89
Query Match:        29.98%            Indels:      23
DB:                  9                Gaps:         5

US-09-735-713a-2 (1-306) x US-09-978-191A-170 (1-1327)

QY  54  SerArgileuGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu  73
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db  534  AGCCGATGTTGGTGAGCTGTCTCCGAGGTGAGTGCCATGGCAGCCAGCCCTC  593

QY  74  LysGlnArgGlnLysHisIleCysGlySerIleValSerProGlnTrpValIleThr  93
      ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db  594  CAGGTTCCGGGTCGACACATCTGTGGGGGCCCTCATCGCTACCGCTGGTGATAACA  653

QY  94  AlaAlaHisCysIleAlaAsnArgAsnIleValSerThrLeu-----AsnValThrAla  111
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db  654  GTTGCCCACTGTTCCAGGAGGACAGCATGGCCTCCACGGTGTGTGGACCGTGTCTCTG  713

QY  112  GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal  131
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db  714  GGCAGGTGTGCAGAACTCGCGCTGGCTGGAGAGGTGTCTTCAAGGTGAGCCGCCCTG  773

QY  132  IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLys  151
      ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db  774  CTCCTGCACCG--TACCACGAGGAGGACGCCATGACTAGACGTGCGCTGTCTGCAG  830

QY  152  MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg  171
      ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db  831  CTCGACACCCCGTGTGTGGCTCGCGCGCGCTGCGCGCGTGTGCTGCGCGCGCTCC  890

QY  172  GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly  191
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db  891  CACTTCTTCGACCCCGCGCTGCACCTGCTGGATTACGGGCTGGGGCGCTTGCAGGGC  950

QY  192  GlyValLeuSerGlnValLeuGlnGluAsnLeuProIleLeuThrTrpGluGluCys  211
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db  951  GSCCCCATCAGCAACGCTCGAGAAAGTGGATGTCAGTTGATCCACAGGACCTGTGC  1010

QY  212  ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly  231

1011  AGCGAGGCTATCGCTTACCAGGTGAGCCCA-----CGCATGCTGTGTGCGCGC  1058
232  PheProAspGlyClyArgAspAlaCysGlnGlyAspSerGlyCysSerLeuMetCysArg  251
1059  TACCGCAAGGCAAGGATGCTGTAGGGTGACTCAGGTGCTCCGCTGGTGTGCAAG  1118
252  AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg  271
1119  GCACTCAGTGGCGCTGGTTCCTGGGGGGCTGGTCACTGGGCGCTGGGCTGTGCGCGC  1178
272  GlyTrpArgAsnValArgLysSerAspGlnGlySerPro-----GlyIlePhe  288
1179  -----CCTAACTACTTCGCGCTCTAC  1199

289  ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly  305
1200  ACCCGCATCAGGTGTGATCAGCTGGATCCAGCAAGTAGTGTGACCTGAGGA  1250

RESULT 13
US-09-978-403A-170
; Sequence 170, Application US/09978403A
; Publication No. US20030050240A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fond, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC17
; CURRENT APPLICATION NUMBER: US/09/978,403A
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
```







Db	714	GGCAAGGTGTGGCAGAACTCGCCCTGGCTGGAGAGGTGTCTTCAAGGTGAGCGCGCTG	773
Qy	132	IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLys	151
Db	774	CTCTGCAACCG--TACCACGAAGAGGACACCATGACTACGACGTGGCCCTGCTGCAG	830
Qy	152	MetaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg	171
Db	831	CTCACCACCCGGTGTGGCTCGGGCGCGGTGCGCCCGCTGCCTGCCCGCGCGCTCC	890
Qy	172	GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly	191
Db	891	CAC TTCCTCAGCGCGCGCTGCACCTGCTGGATTACGGCTGGGGCCCTTCGCGGAGGC	950
Qy	192	GlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTrpGluGluCys	211
Db	951	GGCCCCATCAGCAACCTCTGCAGAAAGTGGATGTGCAGTTGATCCCNACAGGACCTGTC	1010
Qy	212	ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly	231
Db	1011	AGCGAGCGCTATCGCTACCAGGTGACGCCA-----CGCATCTGTGTGCCGCG	1058
Qy	232	PheProAspGlyClyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg	251
Db	1059	TACCGCAGGGCAAGAGGATCCCTCTCAGGGTGACTCAGGTGGTCCGCTGGTGTGAAG	1118
Qy	252	AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg	271
Db	1119	GCAC TACGTGGCGCTGGTCTCTGGCGGGCTGGTCACTGGGCGCTGGGCTGGGCGGG	1178
Qy	272	GlyTrpArgAsnValArgLysSerAspGlnGlySerPro-----GlyIlePhe	288
Db	1179	-----CCTAACTACTCGGCGCTAC	1199
Qy	289	ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly	305
Db	1200	ACCCGGATCAGAGGTGTGATCAGCTGGATCGATCAGCAAGTGGTGCTCCTCAGGA	1250

RESULT 15

```

US-09-978-585A-170
; Sequence 170, Application US/09978585A
; Publication No. US20030049633A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630Pic15

```

Search completed: March 25, 2003, 03:56:34  
Job time : 117 secs



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 25, 2003, 00:14:23 ; Search time 271 Seconds  
(without alignments)  
2542.847 Million cell updates/sec

Title: US-09-735-713A-2  
Perfect score: 1636  
Sequence: 1 MSLSKMLISRNKLLILLGIVE.....IFTDISKVLSWIHEIQTGN 306

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-Q/cgn2\_1/USPTO\_spool/US09735713/runat\_18032003\_124141\_21580/app\_query.fasta\_1.455  
-DB-N\_Geneseq\_101002 -QMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCLALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USPR=US09735713 @CGN\_1\_1\_263 @runat\_18032003\_124141\_21580 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1632	99.8	1262	24	ABK12903
2	1626	99.4	921	22	AAO9328
3	1626	99.4	1568	22	AAO9331
4	1608	98.3	909	22	AAO9329
5	1372.5	83.9	1671	24	ABK31790
6	950	58.1	637	22	AAH98934
7	898	54.9	495	22	AAO9330
8	567.5	34.7	2742	23	AAH88433
9	567.5	34.7	3309	23	AAH87759
10	567.5	34.7	4729	23	AAH72057
11	567.5	34.7	4729	23	AAH74405
12	567.5	34.7	4729	23	AAH77775
13	567.5	34.7	4729	23	AAH84074
14	561.5	34.3	3096	24	ABO86158
15	490.5	30.0	1327	20	AAH34034
16	490.5	30.0	1327	21	AAH78495
17	490.5	30.0	1327	21	AAH58237
18	490.5	30.0	3143	20	AAH34033
19	490.5	30.0	3143	21	AAH78494
20	490.5	30.0	3143	21	AAH58236
21	489.5	29.9	2409	24	ABK31797
22	486.5	29.7	2672	22	AAH13117
23	486.5	29.7	3104	22	AAH13116
24	486	29.7	3736	23	ABL28971
25	476.5	29.1	708	22	AAH13115
26	467	28.5	1656	22	AAH13118
27	467	28.5	2412	20	AAH33949
28	467	28.5	2412	21	AAH78475
29	467	28.5	2412	22	AAH5956
30	467	28.5	2948	24	AAH36726
31	466	28.5	2267	24	ABN95751
32	466	28.5	2267	24	ABL67214
33	466	28.5	2434	23	ABV22421
34	466	28.5	2434	23	ABV28235
35	462	28.2	2413	21	AAH93842
36	462	28.2	2413	22	AAH59907
37	456.5	27.9	1365	24	ABK31787
38	456.5	27.9	1394	21	AAO2323
39	456.5	27.9	1394	24	AAH28649
40	456.5	27.9	2123	24	ABQ55071
41	456.5	27.9	2951	24	AAH36727
42	454.5	27.8	1697	21	AAH77781
43	454.5	27.8	3387	24	ABK31781
44	454.5	27.8	3711	24	ABK86141
45	448	27.4	927	24	ABK16740

ALIGNMENTS

RESULT 1  
ABK12903  
ID ABK12903 standard; cDNA; 1262 BP.  
AC ABK12903;  
XX  
XX  
DT 09-APR-2002 (first entry)  
XX  
DE Human protease PRTS-20 cDNA sequence.  
XX  
KW Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;  
KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;  
KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;  
KW cell proliferative disorder; developmental disorder; epilepsy;  
KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;  
KW reproductive disorder; endometriosis; ss.  
XX  
OS Homo sapiens.

XX Key Location/Qualifiers  
 FH 165..1085  
 FT /\*tag= a  
 FT /product= "Human protease PRTS-20"  
 XX  
 PN WO200198468-A2.  
 XX  
 XX 27-DEC-2001.  
 XX  
 XX 13-JUN-2001; 2001WO-US19178.  
 XX  
 XX 16-JUN-2000; 2000US-212336P.  
 PR 22-JUN-2000; 2000US-213955P.  
 PR 29-JUN-2000; 2000US-215396P.  
 PR 07-JUL-2000; 2000US-216821P.  
 PR 14-JUL-2000; 2000US-218946P.  
 XX  
 XX (INCY-) INCYTE GENOMICS INC.  
 XX  
 XX Yue H, Elliott VS, Gandhi AR, Lal P, Au-young J, Tribouley CM;  
 PI Deleage AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;  
 PI Walla NK, Yao MG, Lu DAM, Patterson C, Tang YT, Walsh RT;  
 PI Azinzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;  
 PI Kallick DA;  
 XX  
 XX WPI; 2002-090437/12.  
 XX P-PSDB: AAU74760.  
 XX  
 XX Twenty one human proteases (referred to as PRTS-1 to PRTS-21), useful  
 PT in the diagnosis, treatment and prevention of gastrointestinal (e.g.  
 PT gastritis), cardiovascular (e.g. atherosclerosis) and cell  
 PT proliferative (e.g. cancer) disorders -  
 XX  
 XX Claim 5; Page 175; 177pp; English.  
 XX  
 XX The present invention relates to twenty one new human proteases,  
 CC referred to as PRTS-1 to PRTS-21. The PRTS polynucleotides and  
 CC polypeptides of the invention are useful in the diagnosis, treatment and  
 CC prevention of gastrointestinal e.g. gastritis, esophageal carcinoma and  
 CC Crohn's disease, cardiovascular e.g. atherosclerosis, hypertension and  
 CC myocardial infarction, autoimmune/inflammatory e.g. acquired  
 CC immunodeficiency syndrome (AIDS), allergies and osteoarthritis, cell  
 CC proliferative e.g. cancer, developmental e.g. Duchenne and Becker  
 CC muscular dystrophy, epithelial e.g. dermatitis, neurological e.g.  
 CC epilepsy and Alzheimer's disease and reproductive e.g. infertility and  
 CC endometriosis disorders. Numerous other examples of each disorder are  
 CC given in the specification. The present nucleic acid sequence encodes  
 CC the human protease PRTS-20 protein of the invention.  
 XX  
 XX SQ Sequence 1262 BP; 354 A; 265 C; 306 G; 337 T; 0 other;  
 XX  
 XX Alignment Scores:  
 XX Pred. No.: 6,74e-149 Length: 1262  
 XX Score: 1632.00 Matches: 305  
 XX Percent Similarity: 99.67% Conservative: 0  
 XX Best Local Similarity: 99.67% Mismatches: 1  
 XX Query Match: 99.76% Indels: 0  
 XX DB: 24 Gaps: 0  
 XX  
 XX US-09-735-713A-2 (1-306) x ABRK12903 (1-1262)  
 XX  
 XX QY 1 MetSerLeuLysMetLeuSerArgAsnLysLeuLeuLeuLeuGlyIleValPhe 20  
 Db ATGAGTCTCAAAATGCTTAAGCAGGAAACAAGCTGATTTTACTAGTAATAGTCTTT 224  
 QY 21 PheGluArgGlyLysSerAlaLeuSerLeuProLysAlaProSerCysGlyGlnSer 40  
 Db TTTGAAAGAGGAGAAATCTCAACTCTTCGCTCCCAAGCTCCAGTTGTGGCAGAGT 284  
 QY 41 LeuValLysValGlnProTyrPheAsnLysPheAsnLysPheSerArgIleLeuGlyLys 60  
 Db CTGGTTAAGGTACAGCTTGGAATATTTTACATTTTCAGTCGCAATCTTGGAGGAAC 344

QY 61 GlnValGluLysGlySerTyrProTyrGlnValSerLeuLysGlnArgGlnLysHisIle 80  
 Db CAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAACAAAGGAGAGCATATT 404  
 QY 81 CysGlyGlySerIleValSerProGlnTyrValIleThrAlaAlaHisCysIleAlaAsn 100  
 Db TGTGGAGGAAGCATCGTCTCACACAGTGGTGTATCAGCGCGCTCACTGCATTGCAAC 464  
 QY 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120  
 Db AGAAACATTGTCTACTTGAATGTACTGTGGAGATGATGACTTAAGCCAGACAGAC 524  
 QY 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140  
 Db CCAGGAGAGCAAACTCTCACTATTGAAACTGTCATCATCATCATCATTTCTCCACCAAG 584  
 QY 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160  
 Db AAACCAATGGACTATGATATTGCCCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGCCAC 644  
 QY 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180  
 Db TTTGTGGGGCCCATATGCTTCCAGAGCTGCGGGAGCAATTTGAGGCTGGTTTATTGT 704  
 QY 181 ThrThrAlaGlyTyrPglyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnGlu 200  
 Db ACAACTGCAGGCTGGGGCCGCTTAAGTGAAGCTGGGCTCTCTCAAGTCTTGCAGGAA 764  
 QY 201 ValAsnLeuProIleLeuThrTyrGluGluCysValAlaLeuLeuThrLeuLysArg 220  
 Db GTGAATCTGCCCTATTTTGACCTGGGAAGAGTGTGGCAGCTCTGTTAACTAAAGAGG 824  
 QY 221 ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys 240  
 Db CCATCAGTGGGAAGACCTTTCTTGACAGGTTTCTCGATGGAGGGAGACCGCATGT 884  
 QY 241 GlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaThrThrLeuAla 260  
 Db CAGGAGAGATTAGGAGGTTCACTCATGTGCCGGAATAAGAAAGGGCGCTGGACTCTGGCT 944  
 QY 261 GlyValThrSerTyrPglyLeuGlyCysGlyArgGlyTyrArgAsnAsnValArgLysSer 280  
 Db GGTGTGACTTCTCGGGTTTGGGCTGTGTCGAGGCTGGAGAACAATGTGAGGAAAGT 1004  
 QY 281 AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTyrPheHisGlu 300  
 Db GATCAAGGATCCCTGGGATCTTCACAGACATTAGTAAAGTCTTCTCTGGATCCACGAA 1064  
 QY 301 HisIleGlnThrGlyAsn 306  
 Db CACATCCAAACTGGTAAAC 1082  
 Db  
 XX RESULT 2  
 XX AAD09328  
 XX ID AAD09328 standard; cDNA; 921 BP.  
 XX AC AAD09328;  
 XX AC AAD09328;  
 XX DT 10-SEP-2001 (first entry)  
 XX  
 XX Human protease homologue cDNA #1.  
 DE Human; protease homologue; novel human protein; NHP; therapy;  
 KW pharmacogenomic application; physiological disorder; ss.  
 XX Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH 1..921  
 FT CDS /\*tag= a  
 FT /product= "Human protease homologue #1"  
 FT unsure 67..69

```

FT      /*tag= b
FT      /note= "Encodes Arg"
FT      68
FT      misc_feature
FT      /*tag= c
FT      /note= "This degenerate base represents a
FT      polymorphic site"
FT      82..84
FT      unsure
FT      /*tag= d
FT      /note= "Encodes Ala"
FT      82
FT      misc_feature
FT      /*tag= e
FT      /note= "This degenerate base represents a
FT      polymorphic site"
FT      82
XX      WO200146407-A1.
XX      28-JUN-2001.
XX      12-DEC-2000; 2000WO-US33738.
XX      23-DEC-1999; 99US-0171566.
XX      (LEXI-) LEXICON GENETICS INC.
XX      PI Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;
XX      WPI: 2001-408641/43.
XX      DR P-PSDB; AAE04733.
XX      PT Polynucleotide encoding novel human protease homologs, useful for
XX      identifying agonist, antagonist or modifiers or for producing
XX      antibodies useful in therapeutic, diagnostic and pharmacogenomic
XX      applications.
XX      PS Claim 1; Page 28; 32pp; English.
XX      CC The present sequence is a cDNA encoding novel human protein (NHP),
XX      CC known as human protease homologue. NHP shares structural similarity
XX      CC with animal proteases, particularly trypsin-like protease such
XX      CC as oviductin, plasminogen activator and human plasma kallikrein
XX      CC precursor. NHP and its cDNA are useful in therapeutic, diagnostic and
XX      CC pharmacogenomic applications. NHP sequences are useful for identifying
XX      CC agonists, antagonists and modulators and also for producing antibodies
XX      CC useful in diagnosis, drug screening, clinical trial monitoring and in
XX      CC treatment of physiological disorders.
XX      SQ Sequence 921 BP; 242 A; 195 C; 240 G; 242 T; 2 other;

Alignment Scores:
Pred. No.: 1.7e-148 Length: 921
Score: 1626.00 Matches: 304
Percent Similarity: 99.35% Conservative: 0
Best Local Similarity: 99.35% Mismatches: 2
Query Match: 99.39% Indels: 0
DB: 22 Gaps: 0

US-09-735-713A-2 (1-306) x RAD09328 (1-921)

QY 1 MetSerLeuLysMetLeuLysSerArgAsnLysLeuLeuLeuLeuGlyLeuValPhe 20
Db 1 ATGAGTCTCAAAATGCTTATAAGCAGGAAAGCTGATTTACTACTAGGAATAGTCTTT 60
QY 21 PheGluArgGlyLysSerAlaLeuSerLeuProLysAlaProSerCysGlyGlnSer 40
Db 61 TTGGAACRAGGTAAATCTCCARCTCTTCCGCTCCCAAGCTCCCAAGTTGTGGGCAGAGT 120
QY 41 LeuValLysValGlnProTrpAsnTyrPheAsnIlePheSerArgIleLeuGlyLys 60
Db 121 CTGGTTAAGGTACAGCCTTGGAAATATTTTACATTTTCAGTCGCATCTTGGAGGAAC 180
QY 61 GlnValGluLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80
Db 181 CAAAGTGAGAGGGTTCCCTATCCCTGGCAGGTATCTCTCAAAACAAGGAGGAGCATATT 240

```

```

QY 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100
Db 241 TGTGGAGGAGCATCGTCTCACCACAGTGGGTGATCACCAGCGGCTCAGTCGATTGCAAC 300
QY 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120
Db 301 AGAACATTTGTCTACTTTGAATGTTACTGCTGGAGAGTATGACTTAAGCCAGACAGAC 360
QY 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140
Db 361 CCAGGAGAGCAAACTCTCACTATTTGAAGTGTCACTATCATCATCATCTCTCCACAAG 420
QY 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160
Db 421 AAACCAATGGACTATGATATTTGCCCTTTGAAGATGGCTGGAGCCTTCCAAATTTGCCAC 480
QY 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180
Db 481 TTTGTGGGCCCATATGCTTCCAGAGCTGGGGAGCAATTTGAGGCTGTTTTATTGT 540
QY 181 ThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValIleuSerGlnValLeuGlnGlu 200
Db 541 ACAACTGCAAGGTGGGGCCCTTAAGTGAAGTGGCTCTCTCAACAAGTCTTTCAGAGAA 600
QY 201 ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg 220
Db 601 GTGATCTCCCTATTTTGACCTGGGAAGAGTGTGGCAGCTCTGTTAACACTTAAGAGG 660
QY 221 ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys 240
Db 661 CCCATCAGTGGGAAGACCTTTCTTTCACAGGTTTTCTGATGAGGAGAGAGCATGT 720
QY 241 GlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAla 260
Db 721 CAGGGAGATTCAGGAGGTTCACTCATGTCCGGAATAAGAAAGGGGCTGGACTCTGGCT 780
QY 261 GlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSer 280
Db 781 GGTGTGACTTCTCTGGGGTTGGGCTGTGGTGGAGGCTGGAGAAACAATGTGAGGAAAGT 840
QY 281 AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu 300
Db 841 GATCAAGGATCCCTGGGATCTTCACAGACATTAGTAAAGTGTCTTCTGATCCAGAA 900
QY 301 HisIleGlnThrGlyAsn 306
Db 901 CACATCCAAACTGGTAAC 918

RESULT 3
AAD09331
ID AAD09331 standard; cDNA; 1568 BP.
XX
AC AAD09331;
XX
DT 10-SEP-2001 (first entry)
XX
DE Human protease homologue cDNA #4.
XX
KW Human; protease homologue; novel human protein; NHP; therapy;
KW pharmacogenomic application; physiological disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200146407-A1.
XX
PD 28-JUN-2001.
XX
PF 12-DEC-2000; 2000WO-US33738.
XX
PR 23-DEC-1999; 99US-0171566.
XX
PA (LEXI-) LEXICON GENETICS INC.

```

XX Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;  
PI WPI; 2001-408641/43.  
XX  
XX Polynucleotide encoding novel human protease homologs, useful for  
PT identifying agonist, antagonist or modifiers or for producing  
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic  
PT applications -  
XX  
XX Disclosure; Page 31; 32pp; English.  
XX  
CC The present sequence is a cDNA encoding novel human protein (NHP),  
CC known as human protease homologue. NHP shares structural similarity  
CC with animal proteases, particularly trypsin-like protease such  
CC as oviductin, plasminogen activator and human plasma kallikrein  
CC precursor. NHP and its cDNA are useful in therapeutic, diagnostic and  
CC pharmacogenomic applications. NHP sequences are useful for identifying  
CC agonists, antagonists and modulators and also for producing antibodies  
CC useful in diagnosis, drug screening, clinical trial monitoring and in  
CC treatment of physiological disorders.  
XX  
SQ Sequence 1568 BP; 452 A; 338 C; 353 G; 423 T; 2 other;

Alignment Scores:  
Pred. No.: 3,43e-148 Length: 1568  
Score: 1626.00 Matches: 304  
Percent Similarity: 99.35% Conservative: 0  
Best Local Similarity: 99.35% Mismatches: 2  
Query Match: 99.39% Indels: 0  
DB: 22 Gaps: 0

US-09-735-713A-2 (1-306) x AAD09331 (1-1568)

Qy 1 MetSerLeuLysMetLeuLysSerArgAsnLysLeuLysLeuLysLeuValPhe 20  
Db 245 ATGAGTCTCAAAATGCTTATAGCAGAACAGCTGATTTTACTAGGAATAGTCTTT 304

Qy 21 PheGluArgGlyLysSerAlaLeuSerLeuProLysAlaProSerCysGlyGlnSer 40  
Db 305 TTTGAACRAGGTAATCTGCARCTTTTCGCTCCCAAGCTCCAGTGTGGCAGAGT 364

Qy 41 LeuValLysValGlnProTrpAsnTrpPheAsnIlePheSerArgIleLeuGlyGlySer 60  
Db 365 CTGTTTAAAGTACAGCCCTTGAATATTTTAAACATTTTTCAGTCGATTTTGGAGGAGC 424

Qy 61 GlnValGluLysGlySerTrpProTrpGlnValSerLeuLysGlnArgGlnLysHistle 80  
Db 425 CAAGTGGAGAAGGGTCTATFCCCTGGCAGGTATCTCTGAACAAAGGCAGACATATT 484

Qy 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100  
Db 485 TGTGGAGGAGCATCGCTCACCACAGTGGGTGATCAGCGGGCTCTACTGCTGCAAC 544

Qy 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTrpAspLeuSerGlnThrAsp 120  
Db 545 AGAAACATTTGTCTACTTTGAATGTTACTTGTCTGAGAGTATGACTTAAAGCCAGACAGAC 604

Qy 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140  
Db 605 CCAGGAGAGCAAACTCTCATATTGAACCTGTCTATCATCATATCCACATTTCTCCACCAAG 664

Qy 141 LysProMetAspTrpAspIleAlaLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160  
Db 665 AAACCATGGACTATGATATTTGCCCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCAC 724

Qy 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180  
Db 725 TTTCTGGGGCCCATATGCTTCCAGAGCTGCGGAGCAATTTTGAAGCTGGTGTATTTTGT 784

Qy 181 ThrThrAlaGlyTrpGlyArgLeuThrGluGlyValLeuSerGlnValLeuGlnGlu 200  
Db 785 ACAACTGCAGGCTGGGGCCGCTTAACTGAAGGTGGCGCTCTCTCACAAGTCTTTCAGGAA 844

Qy 201 ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg 220  
Db 845 GTGAATCTGCTATTTTGCCTGGGAAGAGTGTGTGGCAGCTCTGTTAACACTAAAGAGG 904

Qy 221 ProfileSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys 240  
Db 905 CCCATCAGTGGGAAGACCTTTCTTTTGACAGGTTTCTCTGATGGAGGAGAGCGCATGT 964

Qy 241 GlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAla 260  
Db 965 CAGGAGGATTCAGGAGGTTCTACTGTGCGGGAATAAGAAAGGGCTGGACTTGGCT 1024

Qy 261 GlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnAsnValArgLysSer 280  
Db 1025 GGTCTGACTCTCTGGGTTTGGGCTGTGGTCAGGCTGGAGAACATGTGAGGAAAGT 1084

Qy 281 AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu 300  
Db 1085 GATCAAGGATCCCTGGGATCTTTCACAGACATATTAGTAAAGTCTTCTCTGGATCCACGAA 1144

Qy 301 HistleGlnThrGlyAsn 306  
Db 1145 CACATCCAAACTGGTAAC 1162

RESULT 4  
AAD09329  
ID AAD09329 standard; CDNA; 909 BP.  
XX  
AC AAD09329;  
XX  
DT 10-SEP-2001 (first entry)  
DE Human protease homologue cDNA #2.  
XX  
KW Human; protease homologue; novel human protein; NHP; therapy;  
KW pharmacogenomic application; physiological disorder; ss.  
XX  
XX Homo sapiens.  
PH Key Location/Qualifiers  
CDS 1..909  
FT /\*tag= a  
FT /product= "Human protease homologue #2"  
FT unsure 55..57  
FT /\*tag= b  
FT /note= "Encodes Arg"  
FT misc\_feature 56  
FT /\*tag= c  
FT /note= "This degenerate base represents a  
FT polymorphic site"  
FT unsure 70..72  
FT /\*tag= d  
FT /note= "Encodes Ala"  
FT misc\_feature 70  
FT /\*tag= e  
FT /note= "This degenerate base represents a  
FT polymorphic site"  
XX  
PN WO200146407-A1.  
XX  
PD 28-JUN-2001.  
PF 12-DEC-2000; 2000WO-US33738.  
XX  
PR 23-DEC-1999; 99US-0171566.  
XX (LEXI-) LEXICON GENETICS INC.  
XX Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;  
XX WPI; 2001-408641/43.  
DR P-PSDB; AAE04734.  
DR



CC hypertension, psychotic disorders, neurological disorders  
CC (e.g. Alzheimer's disease, Parkinson's disease) and dyskinesias.  
CC The nucleic acids and polypeptides are also useful for treating viral  
CC infections caused by human immunodeficiency virus (HIV), and non-viral  
CC infections such as ocular disease (e.g. glaucoma) and macular  
CC degeneration. ABK31744-ABK31802 represent DNA sequences encoding for  
CC the novel human proteases of the invention.  
XX

SQ Sequence 1671 BP; 451 A; 369 C; 404 G; 447 T; 0 other;

Alignment Scores:

Pred. No.: 1.54e-123 Length: 1671  
Score: 1372.50 Matches: 266  
Percent Similarity: 93.75% Conservative: 4  
Best Local Similarity: 92.36% Mismatches: 6  
Query Match: 83.89% Indels: 12  
DB: 24 Gaps: 2

US-09-735-713A-2 (1-306) x ABK31790 (1-1671)

Qy 1 MetSerLeuLysMetLeuIleSerArgAsnLysLeuIleLeuLeuGlyIleValPhe 20  
Db 1 ATGAGTCTCAAAATGCTTATAGCAGAACAGCTGATTTTACTACTAGGAATAGTCTTT 60  
Qy 21 PheGluArgGlyLysSerAlaLalaLeuSerLeuProLysAlaProSerCysGlyGlnSer 40  
Db 61 TTGACAGAGGTAATCTGCAATCTTTTCGCTCCCAAGCTCCAGTTGTGGCAGAGT 120  
Qy 41 LeuValLysValGlnProTrpAsnTrpPheAsnIlePheSerArgIleLeuGlyLys 60  
Db 121 CTGGTTAAGTGACAGCTTGGAAATATTTTAACTTTTCAGTGGCATTTTGGAGGAAGC 180  
Qy 61 GlnValGluGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80  
Db 181 CAAAGTGACAGAGGTTCTCTATCCCTGGCAGGTATCTCTGAACAAAGGAGGAGCATATT 240  
Qy 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100  
Db 241 TGTGGAGGAAGCATCTCTCACCACAGTGGTGTATCATCAGCGCGCTCAGTGCATTGCAAC 300  
Qy 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120  
Db 301 AGAAACATTGTGTCTACTTTGAATGTTACTGCTGGAGAGTATGACTTAAAGCCAGACAGAC 360  
Qy 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140  
Db 361 CCAGGAGACCAACTCTACTATTGAACTGTATCATCATCATCATCTTCCACCAAG 420  
Qy 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160  
Db 421 AAACCAATGGACTATGATATTGCCCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGCCAC 480  
Qy 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180  
Db 481 TTGTGGGGCCCATATGCTTCCAGAGCTGCGGGAGCAATTTGAGGCTGGTTTATTTGT 540  
Qy 181 ThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnGlu 200  
Db 541 ACAACTGCAGGCTGGGGCGCTTAAGTGAAGTGGCTCTCTCAAGCTCTTGAGGNA 600  
Qy 201 ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg 220  
Db 601 GTGAATCTGCTTATTTTACCTTGGGAAGAGTGTGTGGCAGCTCTGTTAACACTAAAGAGG 660  
Qy 221 ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyClyArgAspAlaCys 240  
Db 661 CCCATCAGTGGGAAGACCTTTCTTTCACAGGTTTTCCTGTATGGAGGAGACCATGT 720  
Qy 241 GlnGlyAspSerGlyCysSerLeuMetCysArgAsnLysLysGlyAlaValTrp-ThrLeuAl 260  
Db 721 CAGGAGATTACAGAGGTTCACTCATGTGCCGGAATAGAAAGGGCCCTGGGACTCTGGC 780  
Qy 260 aclyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValIleArgLysSe 280

Db 781 TGGT-----CAATTTGG-----AGGCTCAGGTGGGA 807  
Qy 280 rAspGlnGlySerProGlyIle 287  
Db 808 GGATCGCTTGGTCCAGTCCAGGAGTT 829  
RESULT 6  
AAH98934  
ID AAH98934 standard; cDNA; 637 BP.  
XX  
AC AAH98934;  
XX  
DT 12-OCT-2001 (first entry)  
XX  
DE Toad EST-derived coding sequence SEQ ID NO: 791.  
XX  
KW Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
KW diagnostics; forensic test; gene mapping; genetic disorder;  
KW biodiversity; gene therapy; nutrition; ss.  
OS Xenopus laevis.  
XX  
PN WO200154477-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 25-JAN-2001; 2001WO-US02687.  
XX  
PR 25-JAN-2000; 2000US-0491404.  
PR 17-JUL-2000; 2000US-0617746.  
PR 03-AUG-2000; 2000US-0631451.  
PR 15-SEP-2000; 2000US-0663870.  
XX  
PA (HYSEQ-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
PI Cao Y, Drmanac RA, Zhang J, Werhman T;  
XX  
DR WPI; 2001-476164/51.  
DR P-PSDB; AAM24275.  
XX  
PT Isolated polypeptide for treatment of diseases, diagnostics, raising  
PT antibodies and research use -  
XX  
PS Claim 1; Page 673-674; 1275pp; English.  
XX  
CC The present invention provides the protein and coding sequences of novel  
CC proteins from a variety of organisms, including human, dog, cat, horse,  
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)  
CC from the organism of interest. They can be used in diagnostics,  
CC forensics, gene mapping, identification of mutations, to assess  
CC biodiversity and for nutritional purposes. The present sequence is a cDNA  
CC of the invention.  
XX  
SQ Sequence 637 BP; 172 A; 133 C; 141 G; 191 T; 0 other;

Alignment Scores:

Pred. No.: 4.62e-83 Length: 637  
Score: 950.00 Matches: 181  
Percent Similarity: 95.00% Conservative: 0  
Best Local Similarity: 99.45% Mismatches: 1  
Query Match: 58.07% Indels: 0  
DB: 22 Gaps: 0

US-09-735-713A-2 (1-306) x AAH98934 (1-637)

Qy 1 MetSerLeuLysMetLeuIleSerArgAsnLysLeuIleLeuLeuGlyIleValPhe 20  
Db 91 ATGATCTCAAAATGCTTATAGCAGGAACAAGCTGATTTTACTACTAGGAATAGTCTTT 150

QY 21 PheGluArgGlyLysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGlnSer 40  
 DB 151 TTGTACAGAGGTAATCTGCAACTCTTTCGCTCCCAAGCTCCCAAGTGTGGGACAGAGT 210  
 QY 41 LeuValLysValGlnProTrpAsnTrpPheAsnIlePheSerArgIleLeuGlyGlySer 60  
 DB 211 CTGGTTAAGGTACAGCTTGGAAATTAATTTAAACATTTTCAGTCGCATCTCTGGAGGAAGC 270  
 QY 61 GlnValGluLysGlySerTrpProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80  
 DB 271 CAAAGTGGAGAGGTTCCATCTCCCTGGCAGGTATCTCTGAAACAAAGGAGAGCATATT 330  
 QY 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100  
 DB 331 TGTGGAGGAAGCATCTCTCACACAGTGGGTGATCAGCGCGCTCAGTCGATTGCAAC 390  
 QY 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTrpAspLeuSerGlnThrAsp 120  
 DB 391 AGAAACATTGTGTCTACTTTGAATGTTACTGTGGAGAGTATGACTTAAGCCAGACAGAC 450  
 QY 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140  
 DB 451 CCAGGAGAGAACTCTCACTATTTGAACCTGTCATCATCATCATCATCTTCTCCACCAAG 510  
 QY 141 LysProMetAspTrpAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160  
 DB 511 AAACCAATGGACTATGATATGCGCTTTTGAAGATGGCTGGAGCTTCCAAATTTGGCCAC 570  
 QY 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180  
 DB 571 TTTGTGGGCCCCATATGCTTCCAGAGCTGCGGGAGCAATTTGAGCTGGTTTATTGT 630  
 QY 181 ThrThr 182  
 DB 631 ACAACT 636  
 RESULT 7  
 AAD09330  
 ID AAD09330 standard; cDNA; 495 BP.  
 XX  
 AC AAD09330;  
 XX  
 DT 10-SEP-2001 (first entry)  
 XX  
 DE Human protease homologue cDNA #3.  
 XX  
 KW Human; protease homologue; novel human protein; NHP; therapy;  
 KW Pharmacogenomic application; physiological disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..495  
 FT /tag= a  
 FT /product= "Human protease homologue #3"  
 XX  
 FT  
 XX WO200146407-A1.  
 XX  
 PD 28-JUN-2001.  
 XX  
 PF 12-DEC-2000; 2000WO-US33738.  
 XX  
 PR 23-DEC-1999; 99US-0171566.  
 XX  
 XX (LEXI-) LEXICON GENETICS INC.  
 XX  
 PI Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;  
 DR WPI: 2001-408641/43.  
 DR P-PSDB; AAE04735.  
 XX  
 PT Polynucleotide encoding novel human protease homologs, useful for  
 PT Identifying agonist, antagonist or modifiers or for producing

PT antibodies useful in therapeutic, diagnostic and pharmacogenomic  
 PT applications -  
 XX  
 PS Disclosure; Page 30; 32pp; English.  
 XX  
 CC The present sequence is a cDNA encoding novel human protein (NHP),  
 CC known as human protease homologue. NHP shares structural similarity  
 CC with animal proteases, particularly trypsin-like protease such  
 CC as oviductin, plasminogen activator and human plasma kallikrein  
 CC precursor. NHP and its cDNA are useful in therapeutic, diagnostic and  
 CC pharmacogenomic applications. NHP sequences are useful for identifying  
 CC agonists, antagonists and modulators and also for producing antibodies  
 CC useful in diagnosis, drug screening, clinical trial monitoring and in  
 CC treatment of physiological disorders.  
 XX  
 SQ Sequence 495 BP; 117 A; 101 C; 147 G; 130 T; 0 other;

Alignment Scores:  
 Pred. No.: 3.72e-78 Length: 495  
 Score: 898.00 Matches: 164  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 54.89% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-735-713a-2 (1-306) x AAD09330 (1-495)

QY 143 MetAspTrpAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHisPheVal 162  
 DB 1 ATGGACTATGATATTCCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCACTTTGTG 60  
 QY 163 GlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThr 182  
 DB 61 GGCCCATATGCTCTCCAGAGCTGCGGGAGCAATTTGAGCTGGTTTATTGTACAAC 120  
 QY 183 AlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnGluValAsn 202  
 DB 121 GCAGGCTGGGCGGCTTAACCTGAAGGTGGCGTCTCTCACAAAGTCTTGCAGGAAGTGAAT 180  
 QY 203 LeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArgProIle 222  
 DB 181 CTGCCCTATTTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTAAACACTAAAGAGGCCCATC 240  
 QY 223 SerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCysGlnGly 242  
 DB 241 AGTGGGAAGACCTTCTTTCGACAGGTTTCTGTATGGAGGAGAGACCGCATGTCAGGGA 300  
 QY 243 AspSerGlyGlySerLeuMetCysArgAsnLysGlyAlaTrpThrLeuAlaGlyVal 262  
 DB 301 GATTCAGGAGGTTCACTCATGTGCCGGAATAAGAAAGGGCCCTGGACTCTGGCTGGTGTG 360  
 QY 263 ThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnAsnValArgLysSerAspGln 282  
 DB 361 ACTTCTGGGGTTGGGCTGTGCTGAGGCTGGAGAACAAATGTGAGGAAAGTGTATCAA 420  
 QY 283 GlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIle 302  
 DB 421 GGATCCCCCTGGGATCTTCACAGACATTAAGTAAAGTGGTTCCTCGGATCCACGACATC 480  
 QY 303 GlnThrGlyAsn 306  
 DB 481 CAAACTGGTAAAC 492  
 RESULT 8  
 AAS88433  
 ID AAS88433 standard; cDNA; 2742 BP.  
 XX  
 AC AAS88433;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE DNA encoding novel human diagnostic protein #24237.  
 XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX Homo sapiens.  
 XX WO200175067-A2.  
 PN 11-OCT-2001.  
 XX 30-MAR-2001; 2001WO-US08631.  
 XX 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX (HYSE-) HYSEQ INC.  
 XX Drmanac RT, Liu C, Tang YT;  
 XX WPI; 2001-639362/73.  
 DR P-PSDB; ABG24246.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity -  
 XX Claim 1; SEQ ID NO 24237; 103pp; English.  
 XX The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS84197-AAS94564 represent novel human  
 CC diagnostic coding sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 2742 BP; 835 A; 561 C; 664 G; 682 T; 0 other;

#### Alignment Scores:

Pred. No.: 4.35e-45 Length: 2742  
 Score: 567.50 Matches: 114  
 Percent Similarity: 59.57% Conservative: 51  
 Best Local Similarity: 41.16% Mismatches: 101  
 Query Match: 34.69% Indels: 11  
 DB: 23 Caps: 4

US-09-735-713A-2 (1-306) x AAS88433 (1-2742)

Qy 37 CysGlyGlnSerLeuValLysVal-----GlnProTrpAsnTyrPheAsnIlePhe 53  
 Db 25 TGTGGAATTCGATGGTCAACATGAAAGTAAGAACCTGCGGTGGGATCTAGATTCTC 84  
 Qy 54 SerArgIleLeuGlyGlySerGlnValclulysGlySerTyrProTrpGlnValSerLeu 73  
 Db 85 TCTAGAATTAGTAGTTGGAGAAATTCACAGTACTGCATCCATCGCAGGTCTCCCTA 144  
 Qy 74 LysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpValIleThr 93  
 Db 145 AATCAGATGAGCACCATTCTGTGGAGGAGCTTGATTCAAGAGATCGGGTGTGTACA 204

Qy 94 AlaAlaHisCysIle-----AlaAsnArgAsnIleValSerThrLeuAsnValThrAla 111  
 Db 205 GCAGCACACTGCCTGGACAGCCTCAGTGAGAGCAGCTGAAGAATATAAAGTGTGACTTCT 264  
 Qy 112 GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131  
 Db 265 GGGGAGTACAGCCTCTTTTTCAGAGGATAAGCAAGAACAGAAATATTCCTGTCTCAAAAT 324  
 Qy 132 IleIleHisProHisPheSerThrLysProMetAspTyrAspIleAlaLeuLeuLys 151  
 Db 325 ATTACCATCTGTAATACAGCCGCTGAATATATAGTCTCTGATATTCGACTGCTGAT 384  
 Qy 152 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171  
 Db 385 CTAAACACAAAGTCAAGTTTGGAAATCTCTTCAGCCAACTGTCTTCCTGACAGCGAT 444  
 Qy 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly 191  
 Db 445 GATAAAGTTGAACCAAGCAAGAAATCTTTTGTCTTATCCAGTGGATGGGCAAGATTTCCAAACA 504  
 Qy 192 GlyValLeuSerGlnValLeuGlnValAsnLeuProIleLeuThrTrpGluGluCys 211  
 Db 505 TCAGAAATATTCAAATGCTCTACAAAGAAATGGAATCTCCATCATGGATGACAGAGCGTGT 564  
 Qy 212 ValAlaAlaLeuThrLeuLysArgProIleSerGlyLysThrPheLeuLeuCysThrGly 231  
 Db 565 AATACTGTGCTCAAGACGATGAACCTCCCTCCCTGGGAAGACCATCTGTGTGCTGGC 624  
 Qy 232 PheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg 251  
 Db 625 TTCCTGATTGGGAATGGACGCTCCAGGGGACTCTGGAGGACCACTGGTTGTGAGA 684  
 Qy 252 AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg 271  
 Db 685 AGAGGTGGTGAATCTGGATTCTTGTGGGATAACTCTCTGGGTAGCTGGTGTGCTGGA 744  
 Qy 272 Gly-----TrpArgAsnAsnValArgLysSerAspGlnGlySerProGlyIlePhe 288  
 Db 745 GGTTCAGTTCCTCCGTGAAGAACACCATCTGAAGCA-----TCACTTGGCATTTTC 795  
 Qy 289 ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly 305  
 Db 796 TCCAAAGTGTCTGATGATTGATGATTTATCATTCAAAACCTGTTTCACAGGT 846  
 RESULT 9  
 AAS87759/c  
 ID AAS87759 standard; cDNA; 3309 BP.  
 XX AC AAS87759;  
 XX DT 13-FEB-2002 (first entry)  
 XX DE DNA encoding novel human diagnostic protein #23563.  
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 XX food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX OS Homo sapiens.  
 XX PN WO200175067-A2.  
 XX PD 11-OCT-2001.  
 XX PF 30-MAR-2001; 2001WO-US08631.  
 XX PR 31-MAR-2000; 2000US-0540217.  
 XX PR 23-AUG-2000; 2000US-0649167.  
 XX PA (HYSE-) HYSEQ INC.  
 XX PI Drmanac RT, Liu C, Tang YT;  
 XX

DR WPI: 2001-639362/73.  
 DR P-PSDB; ABG23572.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity -  
 XX  
 PS Claim 1; SEQ ID No 23563; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human  
 CC diagnostic coding sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 3309 BP; 789 A; 783 C; 683 G; 1054 T; 0 other;

## Alignment Scores:

Pred. No.:	5,56e-45	Length:	3309
Score:	567.50	Matches:	114
Percent Similarity:	59.57%	Conservative:	51
Best Local Similarity:	41.16%	Mismatches:	101
Query Match:	34.69%	Indels:	11
DB:	23	Gaps:	4

US-09-735-713A-2 (1-306) x AAS87759 (1-3309)

Qy	37	CysGlyGlnSerLeuValVal-----GlnProTrpAsnTyrPheAsnIlePhe	53
Db	3285	TGTGGAAATTCGATGTCACATGAAAGTAAGGAACCTCCCGTGGGATCTAGATTCTTC	3226
Qy	54	SerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu	73
Db	3225	TCTAGATTAGTAGTTCGGAGAAATTCACAGTGNCTGGACATCCATCGCGAGGTCCTCCTA	3166
Qy	74	LysGlnArgGlnLysIleCysGlyGlySerIleValSerProGlnTrpValIleThr	93
Db	3165	AAATTCAGATGAGCACCACCTCTCTGTGGAGGAAGCTTGATTCAGGAAGATCGGGTGTGTACA	3106
Qy	94	AlaAlaHisCysIle-----AlaAsnArgAsnIleValSerThrLeuAsnValThra	111
Db	3105	GCAGCACATTCGCTGGACAGCCCTCAGTGAAGCAGCTGAAGATATAACTGTGACTCTCT	3046
Qy	112	GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal	131
Db	3045	GGGGAGTACAGCTCTTTCAGAGGATAACCAAGACAGATATCTCTGTCTCAAAAATT	2986
Qy	132	IleIleHisProHisSerThrLysSerProMetAspTyrAspIleAlaLeuLeuLys	151
Db	2985	ATTACCCATCCTGAATACACAGCGGTGAATATATGATGCTCTGATATGCACCTCTGTAT	2926
Qy	152	MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg	171
Db	2925	CTAAACACAAAGTCAAGTGTGAAATGCTTTCAGCCAAATCTCTCTCTCTGACAGCGAT	2866
Qy	172	GluGlnPheGluAlaGlyPheIleCysThrThraAlaGlyTrpGlyArgLeuThrGluCly	191

Db	2865	GATAAAGTTGACACAGGAATTCCTTGTCTATCCAGTGGTGGCAAGATTCTTCAAAACA	2806
Qy	192	GlyValLeuSerGlnValLeuGlnValAsnLeuProIleLeuThrTrpGluGluCys	211
Db	2805	TCAGAATATTCAAAATGCTCTACAGAATGGAATGGAATTCCTCCATCATGATGACAGACGCTGT	2746
Qy	212	ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly	231
Db	2745	AATACTGTCTCAAGACGATCAACCTCCCTCCCTGGAGGACCATGCTGTGTCTGCGC	2686
Qy	232	PheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg	251
Db	2685	TTCCCTGATGGGGAATGGAGCGCTCCAGGGGACTCTGGAGGACCACTGGTTGTGTAGA	2626
Qy	252	AsnLysLysGlyAlaThrThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg	271
Db	2625	AGAGGTGGTGAATCTCGATTCTTGTGGATACCTTCTCTGGTAGCTGGTGTGCTGGA	2566
Qy	272	Gly-----TrpArgAsnAsnValArgLysSerAspGlnGlySerProGlyIlePhe	288
Db	2565	GGTTCAGTTCCGTAAAGAACCAACCATGTGAAGCA-----TCACTTGGCATTTTC	2515
Qy	289	ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly	305
Db	2514	TCCAAAGTCTGTGATGTGATGATTTATCATCTCAAAACCTGTTTCACAGGT	2464

RESULT 10  
 AAS72057  
 ID AAS72057 standard; cDNA; 4729 BP.  
 XX  
 AC AAS72057;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE DNA encoding novel human diagnostic protein #7861.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US08631.  
 XX  
 PR 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 XX Drmanac RT, Liu C, Tang YT;  
 PI WPI: 2001-639362/73.  
 DR P-PSDB; ABG07870.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity -  
 XX  
 PS Claim 1; SEQ ID No 7861; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX

SQ Sequence 4729 BP; 1652 A; 1025 C; 960 G; 1092 T; 0 other;

#### Alignment Scores:

Pred. No.: 8.89e-45 Length: 4729  
Score: 567.50 Matches: 114  
Percent Similarity: 59.57% Conservative: 51  
Best Local Similarity: 41.16% Mismatches: 101  
Query Match: 34.69% Indels: 11  
DB: 23 Gaps: 4

US-09-735-713A-2 (1-306) x AAS72057 (1-4729)

QY 37 CysGlyGlnSerLeuValVal-----GlnProTrpAsnTyrPheAsnIlePhe 53  
DB 2542 TGTGGAATTCGCATCGTCAACATGAAAGTAAGAACCTGCCGTGGGATCAGATTCATC 2601  
QY 54 SerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu 73  
DB 2602 TCTAGATTAGTAGTGGAGAAATTCACAGTGCAGTGCATGCCATGCCAGTCTCCCTA 2661  
QY 74 LysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpValIleThr 93  
DB 2662 AATCAGATGACACCATCTCTGTGGAGAGCTTGATTCAGAGAGATCGGGTGTTCACA 2721  
QY 94 AlaAlaHisCysIle-----AlaAsnArgAsnIleValSerThrLeuAsnValThrAla 111  
DB 2722 GCAGCACACTGCTGCAGACCTCAGTCAGAGACGAGTGAAGAATATAACTGTGACTTCT 2781  
QY 112 GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131  
DB 2782 GGGGAGTACAGCTCTTCAGAGGATAAGCAAGACAGATAATATTCCTGCTCAAAAATT 2841  
QY 132 IleIleHisProHisSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLys 151  
DB 2842 ATTACCATCTGATACACAGCCGCTGAATATATGAGTCTCTGATATTGCAGCTGCTGAT 2901  
QY 152 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171  
DB 2902 CTAACAACAAAGTCAAGTTGGAAATGCTGTTACGCCAATCTGTCTTCTGACAGCGAT 2961  
QY 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTyrGlyArgLeuThrGluGly 191  
DB 2962 GATAAAGTTGAACACAGGAATCTTTGCTTATCCAGTGGTGGGCAAGATTTCACAAACA 3021  
QY 192 GlyValLeuSerGlnValGlnGlnValAsnLeuProIleLeuThrTyrPgluGluCys 211  
DB 3022 TCAGATATTCATATGCTCTACAGAATAGGAATTCCTCCATCATGGATGACAGACGCTGT 3081  
QY 212 ValAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly 231  
DB 3082 AATACTGTGCTCAAGACGATGAACCTCCCTCCCTGGGAAGGACCATGCTGTGCTGGC 3141  
QY 232 PheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg 251  
DB 3142 TTCCCTGATTGGGGAATGGACGCTGCCAGGGGGACTCTGGAGGACCATGCTGTGTGTA 3201  
QY 252 AsnLysLysGlyAlaGlyThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg 271  
DB 3202 AGAGTGTGGNATCGATCTTCTGGGAAATACCTTCTGGGTAGCTGGTGTGCTCGA 3261

QY 272 Gly-----TrpArgAsnAsnValArgLysSerAspGlnGlySerProGlyIlePhe 288  
DB 3262 GGTTCAGTCCCGTAAAGAACCAACCATGTAAGGCA-----TCATTTGGCATTTTC 3312  
QY 289 ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly 305  
DB 3313 TCCAAAGTGTCTGAGTTGATGATGATTTATCTACTCAAAACCTGTCACAGGT 3363  
RESULT 11  
AAS74405  
ID AAS74405 standard; cDNA; 4729 BP.  
XX  
AC AAS74405;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #10209.  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
OS Homo sapiens.  
XX  
XX WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US08631.  
XX  
PR 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR WPI; 2001-639362/73.  
DR P-PSDB; ABG10218.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
PS Claim 1; SEQ ID No 10209; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 4729 BP; 1652 A; 1025 C; 960 G; 1092 T; 0 other;

Alignment Scores:

Pred. No.: 8.89e-45 Length: 4729

Score:	567.50	Matches:	114
Percent Similarity:	59.57%	Conservative:	51
Best Local Similarity:	41.16%	Mismatches:	101
Query Match:	34.69%	Indels:	11
DB:	23	Gaps:	4

US-09-735-713A-2 (1-306) x AAS74405 (1-4729)

Qy 37 CysGlyClnSerLeuVallysVal-----GlnProTrpAsnTyrPheAsnIlePhe 53  
Db 2542 TGTGGAATTGCATGTGTCACATGAAACTAAGGAACCTGCCGTGGATCTAGATTCTTC 2601  
Qy 54 SerArgIleLeuGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu 73  
Db 2602 TCTAGAAATTAGTAGTTGGAGAAATTCAACAGTGACTGGACATCCATGGCAGGTCTCCCTA 2661  
Qy 74 LysGlnArgGlnLysIleCysGlyGlySerIleValSerProGlnTrpValIleThr 93  
Db 2662 AAATCAGATCAGCACCACCTCTGTGGAGAACCTTGATTCAGGAAGATCGGGTGTTCACA 2721  
Qy 94 AlaAlaHisCysIle-----AlaAsnArgAsnIleValSerThrLeuAsnValThrAla 111  
Db 2722 GCAGCACACTGCCGTGACAGCCCTCACTGAGAAGCAGCTGGAAGAATATAAAGTGTCACTTCT 2781  
Qy 112 GlyClnTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131  
Db 2782 GGGGAGTACAGCCCTTTTCAGGAAGGATAAGCAAGAACAGAATATTCCTGTCTCAAAAATT 2841  
Qy 132 IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLys 151  
Db 2842 ATTACCCATCTCAATACACAGCCGTGAATATATAGCTCTGATATGTGCACTGCTGTAT 2901  
Qy 152 MetAlaGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171  
Db 2902 CTAAGCAACAAAGTCAAGTTTGGAAATGTGTTCAGCCAAATCTGTCTTCCTGCACAGCGAT 2961  
Qy 172 GluClnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly 191  
Db 2962 GATAAAGTTGAACCCAGGAATCTTTGCTTATCCAGTGGATGGGCAAGATTTCAAAACA 3021  
Qy 192 GlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTrpGluGluCys 211  
Db 3022 TCAGAATATTCAAAATCTCTCAAGAAATGGAACCTCCCATCATGGATGACAGAGCGTGT 3081  
Qy 212 ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly 231  
Db 3082 AATACTGTGCTCAAGAGATGAACCTCCCTCCCTGGGAAGGACCATGCTGTGTGCTGGC 3141  
Qy 232 PheProAspGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg 251  
Db 3142 TTCCCTGATGGGGAATGAGCCCTGCCAGGGGACTCTGGAGGACCACTGGTTGTGTAGA 3201  
Qy 252 AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg 271  
Db 3202 AGAGGTGGTGGAAATCTGGATTCTTGCTGGGAATACTCTCGGTAGCTGGTTGTGCTGGA 3261  
Qy 272 Gly-----TrpArgAsnAsnValArgLysSerAspGlnGlySerProGlyIlePhe 288  
Db 3262 GGTTCAGTTCCCGTAGAAGAACCAACCATGTGAAGGCA-----TCATTGGCATTTTC 3312  
Qy 289 ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly 305  
Db 3313 TCCAAAGTGCTCAGTTGATGGATTATATCATACTCAAAACCTGTTCCACAGGT 3363  
RESULT 12  
AAS78775  
ID AAS78775 standard; cDNA; 4729 BP.  
xx AAS78775;  
AC  
XT  
DT 13-FEB-2002 (first entry)  
xx  
DE DNA encoding novel human diagnostic protein #14579.

XX	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW	food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200175067-A2.
XX	
PD	11-OCT-2001.
XX	
XX	30-MAR-2001; 2001WO-US08631.
PF	
XX	
PR	31-MAR-2000; 2000US-0540217.
XX	
PR	23-AUG-2000; 2000US-0649167.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Drmanac RT, Liu C, Tang YT;
XX	
DR	WPI; 2001-639362/73.
DR	P-PSDB; ABG14588.
XX	
PT	New isolated polynucleotide and encoded polypeptides, useful in
PT	diagnostics, forensics, gene mapping, identification of mutations
PT	responsible for genetic disorders or other traits and to assess
PT	biodiversity
XX	
XX	Claim 1; SEQ ID NO 14579; 103pp; English.
PS	
XX	
CC	The invention relates to isolated polynucleotide (I) and
CC	polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC	polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC	and gene mapping, and in recombinant production of (II). The
CC	polynucleotides are also used in diagnostics as expressed sequence tags
CC	for identifying expressed genes. (I) is useful in gene therapy techniques
CC	to restore normal activity of (II) or to treat disease states involving
CC	(II). (II) is useful for generating antibodies against it, detecting or
CC	quantitating a polypeptide in tissue, as molecular weight markers and as
CC	a food supplement. (II) and its binding partners are useful in medical
CC	imaging of sites expressing (II). (I) and (II) are useful for treating
CC	disorders involving aberrant protein expression or biological activity.
CC	The polypeptide and polynucleotide sequences have applications in
CC	diagnostics, forensics, gene mapping, identification of mutations
CC	responsible for genetic disorders or other traits to assess biodiversity
CC	and to produce other types of data and products dependent on DNA and
CC	amino acid sequences. AAS64197-AAS94564 represent novel human
CC	diagnostic coding sequences of the invention.
CC	Note: The sequence data for this patent did not appear in the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 4729 BP; 1652 A; 1025 C; 960 G; 1092 T; 0 other;
Alignment Scores:	
Pred. No.:	8.89e-45
Score:	567.50
Length:	4729
Matches:	114
Conservative:	51
Mismatches:	101
Indels:	11
Gaps:	4
US-09-735-713A-2 (1-306) x AAS78775 (1-4729)	
Qy	37 CysGlyGlnSerLeuValLysVal-----GlnProTrpAsnTyrPheAsnIlePhe 53
	:    :    :    :
Db	2542 TGTCGAATTCGATGGTCAACATGAAAGTAAGGAAGCTCCGCTGGGATCTAGATTCTTC 2601
Qy	54 SerArgIleuGlyGlySerGlnValGluLysGlySerTyrProTyrGlnValSerLeu 73
	:    :    :    :    :    :    :    :    :
Db	2602 TCTAGAAATTAGTACTTGGAGAAATTCACACAGTGACTGCACATCCATGCAGGTCCTCCTA 2661
Qy	74 LysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpValIleThr 93
	:    :    :    :    :    :    :    :    :    :    :

```
Db 2662 AAATCAGATGACACCACTTCTGTGGAGGAAGCTTTGATTCAAGAAGATCGGTTGTGTACA 2721
Qy 94 AlaAlaHisCysIle-----AlaAsnArgAsnIleValSerThrLeuAsnValThrAla 111
Db 2722 GCAGCACACTGCTGGACAGCTCAGTGTGAGAGCAGCTGAAGAATATTAAGTGTACTGTCT 2781
Qy 112 GlyGluTyAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131
Db 2782 GGGAGTACAGCTCTTTTCAGAGGATAAGCAAGAACAAGATATTCCTGTCTCAAAAATT 2841
Qy 132 IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLys 151
Db 2842 ATTACCCATCTCGAATACACAGCGCTGAATATATAGTCTGATGATTCGATGCTGTAT 2901
Qy 152 MetalGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171
Db 2902 CTAAACACAAAGTCAAGTTTGGAAATCTGTTACGCCAATCTGCTTCTGACAGCGAT 2961
Qy 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly 191
Db 2962 GATNAAGTTGAACCAAGGAATCTTTGCTTATCCAGTGGATGGGCAAGATTTCCAAAACA 3021
Qy 192 GlyValLeuSerGlnValGluGlnValAsnLeuProIleLeuThrTrpGluGluCys 211
Db 3022 TCAGATATTCAAATGCTCTACAGAAATGGAACTTCCATCATGGATGACAGACCGTGT 3081
Qy 212 ValAlaAlaLeuLeuThrLysArgProIleSerGlyLysThrPheLeuCysThrGly 231
Db 3082 AATACTGTGCTCAAGACGATGAACCTCCCTCCCTGGGAAGGACCATGCTGTGCTGCGC 3141
Qy 232 PheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg 251
Db 3142 TTCCCTGATTTGGGGAATGGACGCTCCAGGGGGACTCTGGAGGACCATGCTGTTGTAGA 3201
Qy 252 AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg 271
Db 3202 ACAGGTGTGGAAATCTGGATCTTCTGGGATAAATCTCTGGGTAGTGGTTGCTCGTGA 3261
Qy 272 Gly-----TrpArgAsnAsnValArgLysSerAspGlnGlySerProGlyIlePhe 288
Db 3262 GTTTCAGTCCCGTGAAGAAACAACCATGTGAAGGCA-----TCATTGGCATTTTC 3312
Qy 289 ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly 305
Db 3313 TCCAAAGTGTCTGAGTTGATGATTTTATCACTCAAAACCTGTTCACAGGT 3363

RESULT 13
ID AAS84074
AC AAS84074;
DF 13-FEB-2002 (first entry)
DE DNA encoding novel human diagnostic protein #19878.
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
  food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSEQ-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
```

```
XX WPI; 2001-639362/73.
DR P-PSDB; ABG19887.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
  diagnostics, forensics, gene mapping, identification of mutations
  responsible for genetic disorders or other traits and to assess
  biodiversity.
PS Claim 1; SEQ ID No 19878; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
  polypeptide (II) sequences. (I) is useful as hybridisation probes,
  polymerase chain reaction (PCR) primers, oligomers, and for chromosome
  and gene mapping, and in recombinant production of (II). The
  polynucleotides are also used in diagnostics as expressed sequence tags
  for identifying expressed genes. (I) is useful in gene therapy techniques
  to restore normal activity of (II) or to treat disease states involving
  (II). (II) is useful for generating antibodies against it, detecting or
  quantitating a polypeptide in tissue, as molecular weight markers and as
  a food supplement. (II) and its binding partners are useful in medical
  imaging of sites expressing (II). (I) and (II) are useful for treating
  disorders involving aberrant protein expression or biological activity.
  The polypeptide and polynucleotide sequences have applications in
  diagnostics, forensics, gene mapping, identification of mutations
  responsible for genetic disorders or other traits to assess biodiversity
  and to produce other types of data and products dependent on DNA and
  amino acid sequences. AAS64197-AAS94564 represent novel human
  diagnostic coding sequences of the invention.
XX Note: The sequence data for this patent did not appear in the printed
  specification, but was obtained in electronic format directly from WIPO
  at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 4729 BP; 1652 A; 1025 C; 960 G; 1092 T; 0 other;
```

```
Alignment Scores:
Pred. No.: 8.89e-45 Length: 4729
Score: 567.50 Matches: 114
Percent Similarity: 59.57% Conservative: 51
Best Local Similarity: 41.16% Mismatches: 101
Query Match: 34.69% Indels: 11
DB: 23 Gaps: 4
```

US-09-735-713A-2 (1-306) x AAS84074 (1-4729)

```
Qy 37 CysGlyGlnSerLeuValLysVal-----GlnProTrpAsnTyrPheAsnIlePhe 53
Db 2542 TGTGGAATTCGATGCTCAACATGAAGAAAGTAAGGAACCTGCCGTGGATCTAGATTCTC 2601
Qy 54 SerArgIleLeuGlyGlySerGlnValGluLysGlySerTyrProTrpGlnValSerLeu 73
Db 2602 TCTAGAAATTAGTAGTTGGAGAAATTCACAGTGACTGGACATCCATGCGAGTCTCCCTA 2661
Qy 74 LysGlnArgGlnLysHisIleCysGlyGlySerIleValSerProGlnTrpValIleThr 93
Db 2662 AAATCAGATGAGCACCACCTTCTGTGGAGAGCTTGTATCAAGAAGATCGGTTGTGTACA 2721
Qy 94 AlaAlaHisCysIle-----AlaAsnArgAsnIleValSerThrLeuAsnValThrAla 111
Db 2722 GCAGCACACTGCTGGACAGCTCAGTGAAGAGCAGCTGAAGAATATTAAGTGTACTGTCT 2781
Qy 112 GlyGluTyAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131
Db 2782 GGGAGTACAGCTCTTTTCAGAGGATAAGCAAGAACAAGATATTCCTGTCTCAAAAATT 2841
Qy 132 IleIleHisProHisPheSerThrLysLysProMetAspTyrAspIleAlaLeuLeuLys 151
Db 2842 ATTACCCATCTCGAATACACAGCGCTGAATATATAGTCTGATGATTTGCACTGCTGTAT 2901
Qy 152 MetalGlyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171
Db 2902 CTAAACACAAAGTCAAGTTTGGAAATCTGTTACGCCAATCTGCTTCTGACAGCGAT 2961
```

QY 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrpGlyArgLeuThrGluGly 191  
Db 2962 GATAAAGTTGAACCCAGCAATCTTGTCTATCCAGTGGATGGCGCAAGATTTCACAAACA 3021  
QY 192 GlyValLeuSerGlnValLeuGlnGluValAsnLeuProIleLeuThrTrpGluGluCys 211  
Db 3022 TCAGAAATATTCAAATGCTCTACAGAAATGGAATCTCCCATCATGGATGACAGAGCGTGT 3081  
QY 212 ValAlaAlaLeuLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly 231  
Db 3082 AATACTGCTCAAGACAGTCAACCTCCCTCCCTGGAGGACCATGTGTGTCTGGC 3141  
QY 232 PheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg 251  
Db 3142 TTCCTTGATTGGGAATGGACGCTGCCAGGGGCACTCTGGAGGACCACTGGTTGTAGA 3201  
QY 252 AsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSerTrpGlyLeuGlyCysGlyArg 271  
Db 3202 AGAGTGGTGAATCTGGATCTTGTGGGATAACTTCCTGGTAGCTGGTCTGCTGGA 3261  
QY 272 Gly-----TrpArgAsnAsnValArgLysSerAspGlnGlySerProGlyIlePhe 288  
Db 3262 GGTTCAGTCCCTAGAACACACCATGTCAAGGCA-----TCACTTGGCATTTTC 3312  
QY 289 ThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThrGly 305  
Db 3313 TCCAAAGTGTCTGAGTTGATGATTTATCACTCAAAACCTGTTCACAGT 3363  
RESULT 14  
ABQ86158  
ID ABQ86158 standard; DNA; 3096 BP.  
XX AC ABQ86158;  
XX DT 10-SEP-2002 (first entry)  
XX DE Novel human gene. SEQ ID 29.  
XX KW Human; cytostatic; vulnary; antiarteriosclerotic; antiparkinsonian;  
KW nootropic; neuroprotective; immunosuppressive; haemostatic;  
KW antiinflammatory; cardiant; antiulcer; virucide; antithyroid;  
KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;  
KW wound healing disorders; atherosclerosis; Parkinson's disease;  
KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;  
KW inflammation; neoplastic disease; nervous system disorder;  
KW cardiovascular disorders; pancreatitis; respiratory disorder;  
KW hyperproliferation; systemic autoimmune disease; hyper-immunity;  
KW developmental abnormality; gastrointestinal ulceration; neuropathy;  
KW haematological disease; metabolic disease; sperm dysfunction;  
KW thyroid disorder; hypothyroidism; brain damage; colitis;  
KW cone photo- transduction deficiency; neurological disease; stroke;  
KW anglogenesis; ovulation disorder; spinal cord; thyroid gland; heart;  
KW trachea; thymus; lymph node; muscular system; obesity; anorexia;  
KW growth abnormality; precocious puberty; gene; ss.  
XX OS Homo sapiens.  
XX PN WO200250105-A1.  
XX PD 27-JUN-2002.  
XX PF 17-DEC-2001; 2001WO-US49232.  
XX PR 19-DEC-2000; 2000US-256710P.  
XX PR 20-DEC-2000; 2000US-257048P.  
XX PR 09-JAN-2001; 2001US-260482P.  
XX PR 30-JAN-2001; 2001US-264922P.  
XX PR 06-FEB-2001; 2001US-266797P.  
XX PR 19-MAR-2001; 2001US-276988P.  
XX PR 04-APR-2001; 2001US-281535P.  
XX PR 08-MAY-2001; 2001US-289622P.  
XX PA (SMIK ) SMITHKLINE BEECHAM CORP.

PA (SMIK ) SMITHKLINE BEECHAM PLC.  
XX (GLAX ) GLAXO GROUP LTD.  
PI Agarwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y;  
PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;  
XX WPI; 2002-508784/54.  
DR P-PSDB; ABP60993.  
XX Secreted proteins and polynucleotides useful as vaccines for preventing  
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,  
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune  
PT disorder -  
XX Claim 2(a); Page 245-246; 335pp; English.  
XX The invention relates to an isolated polypeptide with signal sequences  
CC which allow it to be secreted extracellularly or membrane associated.  
CC The activity of polypeptides of the invention may be described as,  
CC cytostatic, vulnary, antiarteriosclerotic, antiparkinsonian, nootropic,  
CC neuroprotective, immunosuppressive, haemostatic, antiinflammatory,  
CC cardiant, antiulcer, virucide, antithyroid, cerebroprotective, anorectic,  
CC and metabolic. Polypeptides and polynucleotides of the invention are  
CC useful in the treatment, or as a vaccine in the prevention of, cancer,  
CC wound healing disorders, infection, atherosclerosis, Parkinson's disease,  
CC and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,  
CC inflammation, neoplastic diseases, nervous system related disorders and  
CC cardiovascular disorders, pancreatitis, respiratory disorder,  
CC hyperproliferation, systemic autoimmune disease, hyper-immunity,  
CC developmental abnormality, gastrointestinal ulceration, neuropathy,  
CC haematological diseases, metabolic diseases, sperm dysfunction, thyroid  
CC disorders e.g. hypothyroidism, brain damages, colitis, cone photo-  
CC transduction deficiency, neurological diseases, stroke, anglogenesis,  
CC ovulation disorders, diseases in the spinal cord, thyroid gland, heart,  
CC trachea, thymus, lymph node and muscular system, obesity, anorexia,  
CC growth abnormalities, and alleviation of precocious puberty. The  
CC sequences given in records ABQ86130-ABQ86184 represent novel human cDNA's  
CC of the invention.  
XX SQ Sequence 3096 BP; 890 A; 619 C; 742 G; 845 T; 0 other;  
Alignment Scores:  
Pred. No.: 1.95e-44 Length: 3096  
Score: 561.50 Matches: 113  
Percent Similarity: 59.42% Conservative: 51  
Best Local Similarity: 40.94% Mismatches: 101  
Query Match: 34.32% Indels: 11  
DB: 24 Gaps: 4  
US-09-735-713A-2 (1-306) x ABQ86158 (1-3096)  
QY 37 CysGlyGlnSerLeuValLysVal-----GlnProTrpAsnTrpPheAsnIlePhe 53  
Db 25 TGTGGAAATTCGCATGTTCACATCAAGAAAGTAAGAACCTGCCCTGGGATCTAGATCTTC 84  
QY 54 SerArgIleLeuGlyGlySerGlnValGluLysGlySerTrpProTrpGlnValSerLeu 73  
Db 85 TCTAGATTAGTAGTTGGAGAAATTCACAGTCACATCGGACATCCATGGCAGGTCTCCCTA 144  
QY 74 LysGlnArgGlnLysHisIleCysGlySerIleValSerProGlnTrpValIleThr 93  
Db 145 AAATCAGATGAGCACCACCTCTCTGGGAGGAAGCTTGATTCAAGAAGATCGGTTGTATCA 204  
QY 94 AlaAlaHisCysIle-----AlaAsnArgAsnIleValSerThrLeuAsnValThrAla 111  
Db 205 GCAGCACACTGCTGGACAGCCCTCAGTGAAGACACGTCAAGAAATATAACTGTGACTTCT 264  
QY 112 GlyGluTyrAspLeuSerGlnThrAspProGlyGluGlnThrLeuThrIleGluThrVal 131  
Db 265 GGGAGTACAGCCTCTTTTCAAGAGGATACGAAGACAGAAATATCTCTGTCTCAAAAT 324  
QY 132 IleIleHisProHisPheSerThrLysLysProMetAspTrpAspIleAlaLeuLeuLys 151  
XX PA (SMIK ) SMITHKLINE BEECHAM CORP.

```
Db 325 ATTACCACTCTGTAATCAACAGCGGTGAATATATATGAGTCTCTGATATTGCACACTGCTGTAT 384
Qy 152 MetalaclyAlaPheGlnPheGlyHisPheValGlyProIleCysLeuProGluLeuArg 171
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 385 CTAACAACAAAGTCAAGTTTGGAAATGCTGTTACAGCAATCTGTTCTTCGTGACAGCGAT 444
Qy 172 GluGlnPheGluAlaGlyPheIleCysThrThrAlaGlyTrrPglyArgLeuThrGluGly 191
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 445 GATAAAGTTGAACACAGGAAATCTTGTCTTATCCAGTGGATGGGCAAGATTTCCAAACA 504
Qy 192 GlyValLeuSerGlnValGlnGluValAsnLeuProIleLeuThrTrrPgluGluCys 211
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 505 TCAGAAATATTCAAATGCTCTACAGAAATGGAATCTCCATCATCGGATGACAGACGCGGT 564
Qy 212 ValAlaAlaLeuThrLeuLysArgProIleSerGlyLysThrPheLeuCysThrGly 231
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 565 AATACTGTGCTCAAGAGATCAACCTCCCTCCCTGGGAAGGACCATGCTGTGTCTGGC 624
Qy 232 PheProAspGlyGlyArgAspAlaCysGlnGlyAspSerGlyGlySerLeuMetCysArg 251
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 625 TTCCCTGATTGGGAATGAGCGCTCCAGCGGGACTCTGGAGGACCATGCTGTGTGTAGA 684
Qy 252 AsnLysLysGlyAlaTrrPthrLeuAlaGlyValThrSerTrrPglyLeuGlyCysGlyArg 271
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 685 AGAGTGTGTGAATCTGGATTCTTCTGGGATACTTCTCGGTAGCTGGTGTGTGGA 744
Qy 272 Gly-----TrrPArgAsnValArgLysSerAspGlnGlySerProGlyIlePhe 288
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 745 GGTTCAGTTCCTCGTAAGAAACCAACCACTGTAAGGCA-----TCACTTGGCATTTTC 795
RESULT 15
AAZ34034
ID AAZ34034 standard; DNA; 1327 BP.
XX AC AAZ34034;
XX XT
XX DT 07-DEC-1999 (first entry)
XX DE Human EST DNA35597.
XX DE
XX KW Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
XX KW probe; blood coagulation disorder; cancer; cellular adhesion disorder;
XX KW secreted protein; transmembrane protein; ss.
XX OS Homo sapiens.
XX PN
XX FN WO9946281-A2.
XX XX
XX PD 16-SEP-1999.
XX PF
XX FF 08-MAR-1999; 99WO-US05028.
XX PR 10-MAR-1998; 98US-0077450.
XX PR 11-MAR-1998; 98US-0077632.
XX PR 11-MAR-1998; 98US-0077641.
XX PR 11-MAR-1998; 98US-0077649.
XX PR 12-MAR-1998; 98US-0077791.
XX PR 13-MAR-1998; 98US-0078004.
XX PR 17-MAR-1998; 98US-0040220.
XX PR 20-MAR-1998; 98US-0078896.
XX PR 20-MAR-1998; 98US-0078910.
XX PR 20-MAR-1998; 98US-0078936.
XX PR 20-MAR-1998; 98US-0078939.
XX PR 25-MAR-1998; 98US-0079294.
XX PR 26-MAR-1998; 98US-0079656.
XX PR 26-MAR-1998; 98US-0079663.
XX PR 27-MAR-1998; 98US-0079664.
XX PR 27-MAR-1998; 98US-0079689.
XX PR 27-MAR-1998; 98US-0079728.
XX PR 27-MAR-1998; 98US-0079786.
XX PR 30-MAR-1998; 98US-0079920.
XX PR 31-MAR-1998; 98US-0080105.
XX PR 31-MAR-1998; 98US-0080107.
XX PR 31-MAR-1998; 98US-0080185.
XX PR 31-MAR-1998; 98US-0080194.
XX PR 01-APR-1998; 98US-0080327.
XX PR 01-APR-1998; 98US-0080338.
XX PR 01-APR-1998; 98US-0080333.
XX PR 01-APR-1998; 98US-0080334.
XX PR 08-APR-1998; 98US-0081049.
XX PR 08-APR-1998; 98US-0081070.
XX PR 08-APR-1998; 98US-0081071.
XX PR 09-APR-1998; 98US-0081195.
XX PR 09-APR-1998; 98US-0081203.
XX PR 09-APR-1998; 98US-0081229.
XX PR 15-APR-1998; 98US-0081817.
XX PR 15-APR-1998; 98US-0081838.
XX PR 15-APR-1998; 98US-0081952.
XX PR 15-APR-1998; 98US-0081955.
XX PR 21-APR-1998; 98US-0082568.
XX PR 21-APR-1998; 98US-0082569.
XX PR 22-APR-1998; 98US-0082700.
XX PR 22-APR-1998; 98US-0082704.
XX PR 23-APR-1998; 98US-0082804.
XX PR 23-APR-1998; 98US-0082767.
XX PR 23-APR-1998; 98US-0082796.
XX PR 28-APR-1998; 98US-0083332.
XX PR 28-APR-1998; 98US-0083336.
XX PR 29-APR-1998; 98US-0083392.
XX PR 29-APR-1998; 98US-0083495.
XX PR 29-APR-1998; 98US-0083496.
XX PR 29-APR-1998; 98US-0083499.
XX PR 29-APR-1998; 98US-0083500.
XX PR 29-APR-1998; 98US-0083545.
XX PR 29-APR-1998; 98US-0083554.
XX PR 29-APR-1998; 98US-0083558.
XX PR 29-APR-1998; 98US-0083559.
XX PR 30-APR-1998; 98US-0083742.
XX PR 05-MAY-1998; 98US-0084366.
XX PR 06-MAY-1998; 98US-0084414.
XX PR 06-MAY-1998; 98US-0084441.
XX PR 07-MAY-1998; 98US-0084598.
XX PR 07-MAY-1998; 98US-0084600.
XX PR 07-MAY-1998; 98US-0084627.
XX PR 07-MAY-1998; 98US-0084637.
XX PR 07-MAY-1998; 98US-0084639.
XX PR 07-MAY-1998; 98US-0084640.
XX PR 07-MAY-1998; 98US-0084643.
XX PR 13-MAY-1998; 98US-0085323.
XX PR 13-MAY-1998; 98US-0085338.
XX PR 13-MAY-1998; 98US-0085339.
XX PR 15-MAY-1998; 98US-0085573.
XX PR 15-MAY-1998; 98US-0085579.
XX PR 15-MAY-1998; 98US-0085580.
XX PR 15-MAY-1998; 98US-0085582.
XX PR 15-MAY-1998; 98US-0085689.
XX PR 15-MAY-1998; 98US-0085697.
XX PR 15-MAY-1998; 98US-0085700.
XX PR 18-MAY-1998; 98US-0085704.
XX PR 18-MAY-1998; 98US-0086023.
XX PR 22-MAY-1998; 98US-0086392.
XX PR 22-MAY-1998; 98US-0086414.
XX PR 22-MAY-1998; 98US-0086430.
XX PR 22-MAY-1998; 98US-0086486.
XX PR 28-MAY-1998; 98US-0087098.
XX PR 28-MAY-1998; 98US-0087106.
XX PR 30-JUL-1998; 98US-0087208.
XX PR 11-SEP-1998; 98US-0094651.
XX PR 98US-0100038.
XX PA (GETH ) GENENTECH INC.
```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model  
Run on: March 25, 2003, 06:30:40 ; Search time 20.5 seconds  
(without alignments)  
3726.800 Million cell updates/sec

Title: US-09-735-713A-1  
Perfect score: 1668  
Sequence: 1 atgagctcaaatgttat.....acatccaaactgtaactaa 921

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_n2p.model -DEV=xlh  
-Q/cgn2\_1/USPTO\_spool/US09735713/runat\_18032003\_124354\_23119/app\_query.fasta\_1.1095  
-DB=SwissProt\_40 -QFMT=fastan -SUFFIX=rsf -MINMATCH=0.1 -LOOPFCI=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09735713 @CGN\_1\_1\_11 @runat\_18032003\_124354\_23119 -NCPU=6 -ICPU=3  
-NO\_XLPAY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	486.5	29.2	786	1 STUB_DROME	Q05319 drosophila
2	466	27.9	638	1 KAL_HUMAN	P03952 homo sapien
3	464	27.8	638	1 KAL_MOUSE	P26262 mus musculu
4	456.5	27.4	454	1 TMS3_HUMAN	P57727 homo sapien
5	448	26.9	342	1 PSS8_RAT	Q9es87 rattus norv
6	446	26.7	342	1 PSS8_MOUSE	Q9esd1 mus musculu
7	446	26.7	638	1 KAL_RAT	PI4272 rattus norv
8	442.5	26.5	1035	1 ENTK_BOVIN	P98072 bos taurus
9	441	26.4	343	1 PSS8_HUMAN	Q16651 homo sapien
10	440.5	26.4	1034	1 ENTK_PIG	P98074 sus scrofa
11	440	26.4	416	1 HEP5_MOUSE	Q35453 mus musculu
12	440	26.4	417	1 HEP5_HUMAN	P05981 homo sapien
13	438	26.3	416	1 HEP5_RAT	Q05511 rattus norv
14	438	26.3	457	1 TMS3_HUMAN	Q9h3s3 homo sapien
15	433	26.0	593	1 FA12_BOVIN	P98140 bos taurus
16	431.5	25.9	1019	1 FA12_HUMAN	P98073 homo sapien
17	427	25.6	311	1 TRY3_MOUSE	Q9qul7 mus musculu
18	426	25.5	855	1 ST14_MOUSE	P56677 mus musculu

19	425.5	25.5	290	1 MPN_HUMAN	Q9bqr3 homo sapien
20	425	25.5	1069	1 ENTK_MOUSE	P97435 mus musculu
21	423	25.4	455	1 TMS3_MOUSE	Q9er04 mus musculu
22	421	25.2	317	1 BSS4_HUMAN	Q9gzn4 homo sapien
23	421	25.2	321	1 TRY3_HUMAN	Q9nrr2 homo sapien
24	419.5	25.1	237	1 TRY3_ASTFL	P00765 astacus flu
25	419	25.1	264	1 CTRL_HUMAN	P40313 homo sapien
26	419	25.1	273	1 TRY3_SHEEP	Q9xsm2 ovis aries
27	415.5	24.9	275	1 TRY3_ANOGA	P35037 anopheles g
28	415	24.9	306	1 BSS4_MOUSE	Q9er10 mus musculu
29	415	24.9	625	1 FA11_HUMAN	P03951 homo sapien
30	414.5	24.9	615	1 FA12_HUMAN	P00748 homo sapien
31	413.5	24.8	263	1 CTR2_CANFA	P04813 canis famil
32	413.5	24.8	274	1 TRY1_HUMAN	P35035 anopheles g
33	413	24.8	437	1 TMS4_HUMAN	Q9nrs4 homo sapien
34	409	24.5	263	1 CTRB_HUMAN	P17538 homo sapien
35	409	24.5	490	1 TMS2_MOUSE	Q9j1q8 mus musculu
36	408.5	24.5	415	1 ACRO_PIG	P08001 sus scrofa
37	408	24.5	275	1 TRY4_ANOGA	P35038 anopheles g
38	407.5	24.4	459	1 PRTC_PIG	Q9g1p2 sus scrofa
39	407	24.4	275	1 TRY3_PIG	Q9n2d1 sus scrofa
40	406	24.3	254	1 TRY3_AEDAE	P29786 aedes aegyp
41	406	24.3	492	1 TMS3_HUMAN	O15393 homo sapien
42	405.5	24.3	603	1 FA12_CAVPO	Q04962 cavia porce
43	404	24.2	407	1 FA7_BOVIN	P22457 bos taurus
44	403	24.2	314	1 TEST_HUMAN	Q9y6m0 homo sapien
45	403	24.2	324	1 TEST_MOUSE	Q9jh17 mus musculu

ALIGNMENTS

RESULT 1  
STUB\_DROME  
ID STUB\_DROME STANDARD; PRT; 786 AA.  
AC Q05319;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Serine proteinase stubble (EC 3.4.21.-) (stubble-stubblid protein).  
GN SB OR SB-SB.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Oregon-R;  
RX MEDLINE=93281671; PubMed=7685111;  
RA Appel L.F., Prout M., Abu-Shumays R., Hammonds A., Garbe J.C.,  
RA Fristrom D., Fristrom J.;  
RT "The Drosophila Stubble-stubblid gene encodes an apparent  
RT transmembrane serine protease required for epithelial  
RT morphogenesis.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:4937-4941(1993).  
CC -I- FUNCTION: HORMONE DEPENDENT PROTEASE REQUIRED FOR EPITHELIAL  
CC MORPHOGENESIS. HAS A DUAL FUNCTION, DETACHES IMAGINAL DISC CELLS  
CC FORM EXTRACELLULAR MATRICES THROUGH ITS EXTRACELLULAR PROTEOLYTIC  
CC DOMAIN AND TRANSMITS AN OUTSIDE-TO-INSIDE SIGNAL TO ITS  
CC INTRACELLULAR DOMAIN TO MODIFY THE CYTOSKELETON DURING  
CC MORPHOGENESIS. MAY BE ABLE TO ACTIVATE ITSELF.  
CC -I- SUBCELLULAR LOCATION: Type II membrane protein.  
CC -I- INDUCTION: BY 20-HYDROXYECDYSONE (20HE).  
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC -I- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-24 IS THE INITIATOR.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).

CC		EMBL; L11451; AAA28918.1; -	
DR	HSP; P00763; IDPO.		
DR	MEROPS; S01.225; -		
DR	FlyBase; FBgn0003319; Sb.		
DR	InterPro; IPR001314; Chymotrypsin.		
DR	InterPro; IPR001294; Ser_protease_Try.		
DR	Pfam; PF00089; trypsin; 1.		
DR	PRINTS; PR00722; CHYMOTRYPSIN.		
DR	SMART; SM00020; Tryp_SPC; 1.		
DR	PROSITE; PS50240; TRYPSIN_DOM; 1.		
DR	PROSITE; PS00134; TRYPSIN_HIS; FALSE_NEG.		
DR	PROSITE; PS00135; TRYPSIN_SER; 1.		
KW	Hydrolase; Serine protease; Transmembrane; Glycoprotein; Zymogen;		
KW	Signal-anchor.		
FT	CHAIN	1	542
FT	CHAIN	543	786
FT	DOMAIN	1	58
FT	TRANSMEM	59	80
FT	DOMAIN	81	786
FT	DOMAIN	543	786
FT	DOMAIN	267	276
FT	DOMAIN	287	298
FT	DOMAIN	391	478
FT	DOMAIN	412	422
FT	DOMAIN	471	478
FT	ACT_SITE	589	599
FT	ACT_SITE	639	639
FT	ACT_SITE	737	737
FT	DISULFID	531	659
FT	DISULFID	574	590
FT	DISULFID	703	722
FT	DISULFID	733	762
FT	CARBOHYD	177	177
FT	CARBOHYD	671	671
SQ	SEQUENCE	786 AA; 85010 MW; CE3E755760B9DEAD CRC64;	
Alignment Scores:			
Pred. No.:	1.54e-37	Length:	786
Score:	486.50	Matches:	107
Percent Similarity:	52.13%	Conservative:	60
Best Local Similarity:	33.44%	Mismatches:	102
Query Match:	29.17%	Indels:	51
DB:	1	Gaps:	11
US-09-735-713A-1 (1-921) x STUB_DROME (1-786)			
QY	25	AGGAACAAGCTGATTTTACTACTAGATAGTCTTTTTTGAACRAGGTPAAATCTGCARCT	84
Db	493	ArgThrProValLeuAlaThrSerGlyIle-----GluThrAsnGluIleSerAsp	509
QY	85	CTTTGCTGCCCAAGCTCCCAAGTGTGGCGAG-----	117
Db	510	SerSerIleProAspAlaGlyAlaLeuGlyArgValLysThrIleSerAlaAlaArgSer	529
QY	118	-----AGTCTGTTAAGGTACAGCCTTGGGAATTATTTTAACATTTTCAGT	162
Db	530	GluCysGlyValProThrLeuAlaArgProGlu-----	541
QY	163	CGCATTTCTGGAGGAGCCAGTGGAGAGGTTCTATCCCTGCAGGTATCTCTGAAA	222
Db	542	ArgIleValGlyGlyLysSerAlaAlaPheGlyArgTrpProTrpGlnValSerValArg	561
QY	223	CAA-----AGGCAGAAGCATATTTGTGGAGAACATCGTCTCACCA	264
Db	562	ArgThrSerPheGlyPheSerSerThrHisArgCysGlyGlyAlaLeuIleAsnGlu	581
QY	265	CAGTGGGTGATCAGCGCGCTCACTGCATTGCCAACAGAAACATGTGTCTACTTTGAAT	324
Db	582	AsnTrpIleAlaThrAlaGlyHisCysValAspAspLeu---LeuIleSerGlnIleArg	600
QY	325	GTTACTGCTGGAGATATGACTTATGACGACAGACCCAGAGACAACTCTCACTATT	384



```

Db 548 CysGlnLysArgTyrGlnAspTyrLys-----IleThrGlnArgMetValCysAla 564
QY 691 GGTTCCTGATGAGGAGAGAGCGATGTCAGGAGATTCAGGAGGTTCCATCATGTGC 750
Db 565 GlyTyrLysGluGlyLysAspAlaCysLysGlyAspSerGlyGlyProLeuValCys 584
QY 751 CGAATAAGAAAGGGCGCTGGACTCTGGCTGGTGTGCTGACTTCCCTGGGGTTGGCTGTGGT 810
Db 585 LysHis---AsnGlyMetTpaArgLeuValGlyIleThrSerTrpGlyGluGlyCysAla 603
QY 811 CGAGGCTGGAGAACATGTCAGGAAAGTGCATCAGGATCCCTGGATCTTCACAGAC 870
Db 604 Arg-----ArgGluGlnProGlyValTyrThrLys 613
QY 871 ATTAGTAAGACTGCTTCTGCTGATCCAGCAACACATCAAACTGGTAAC 918
Db 614 ValAlaGluTyrMetAspTrpIleLeuGluLysThrGlnSerSerAsp 629

RESULT 3
KAL_MOUSE
ID KAL_MOUSE STANDARD; PRT: 638 AA.
AC P26262;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Plasma kallikrein precursor (EC 3.4.21.34) (Plasma prekallikrein)
DE (kininogenin) (Fletcher factor).
DE KLK1 OR KLK3 OR PK.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=BALB/C; TISSUE=Liver;
RX MEDLINE=91090844; PubMed=2264928;
RA Seiden N.G., Sawyer N., Hamelin J., Mion P., Beaubien G.,
RA Brachpaga L., Rochmont J., Mbikay M., Chretien M.;
RT "Mouse plasma kallikrein: cDNA structure, enzyme characterization,
RT and comparison of protein and mRNA levels among species.";
RL DNA Cell Biol. 9:737-748(1990).
CC -!- FUNCTION: THE ENZYME CLEAVES LYS-ARG AND ARG-SER BONDS. IT
CC ACTIVATES, IN A RECIPROCAL REACTION, FACTOR XII AFTER ITS BINDING
CC TO A NEGATIVELY CHARGED SURFACE. IT ALSO RELEASES BRADYKININ FROM
CC HMW KININOGEN AND MAY ALSO PLAY A ROLE IN THE RENIN-ANGIOTENSIN
CC SYSTEM BY CONVERTING PRORENIN INTO RENIN.
CC -!- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Xaa and Lys-|-Xaa
CC bonds, including Lys-|-Arg and Arg-|-Ser bonds in (human)
CC kininogen to release bradykinin.
CC -!- SUBUNIT: THE ZYMOMEN IS ACTIVATED BY FACTOR XIIA, WHICH CLEAVES
CC THE MOLECULE INTO A LIGHT CHAIN, WHICH CONTAINS THE ACTIVE SITE,
CC AND A HEAVY CHAIN, WHICH ASSOCIATES WITH HMW KININOGEN. THESE
CC CHAINS ARE LINKED BY ONE OR MORE DISULFIDE BONDS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMA KALLIKREIN
CC SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 4 APPLE DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M58588; AAA63393.1; -.
CC PIR: A36557; KMSPL.
CC HSP: P00750; IRTF.
CC MEROPS: S01.212; -.
CC MGD: MGI:102849; Klk3.
CC InterPro: IPR000177; Apple.
CC InterPro: IPR001314; Chymotrypsin.

```

```

DR InterPro: IPR003014; PAN.
DR InterPro: IPR001254; Ser_protease_Try.
DR Pfam: PF00024; PAN; 4.
DR Pfam: PF00089; trypsin; 1.
DR PRINTS: PR00005; APPLIEDOMAIN.
DR PRINTS: PR00722; CHYMOTRYPSIN.
DR SMART: SM00223; APPLE; 4.
DR SMART: SM00020; Tryp_SPC; 1.
DR PROSITE: PS00495; APPLE; 4.
DR PROSITE: PS00240; TRYPSIN_DOM; 1.
DR PROSITE: PS00134; TRYPSIN_HIS; 1.
DR PROSITE: PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease; Glycoprotein; Plasma; Zymogen; Signal;
KW Fibrinolysis; Blood coagulation; Inflammatory response; Liver;
KW Repeat.
FT SIGNAL 1 19 PLASMA KALLIKREIN HEAVY CHAIN.
FT CHAIN 20 390 PLASMA KALLIKREIN LIGHT CHAIN.
FT DOMAIN 391 638 APPLE 1.
FT DOMAIN 20 105 APPLE 2.
FT DOMAIN 110 195 APPLE 3.
FT DOMAIN 200 285 APPLE 4.
FT DOMAIN 291 376 SERINE PROTEASE.
FT DOMAIN 389 621 N-LINKED (GLCNAC. .) (PROBABLE).
FT CARBOHYD 127 127 N-LINKED (GLCNAC. .) (PROBABLE).
FT CARBOHYD 308 308 N-LINKED (GLCNAC. .) (PROBABLE).
FT CARBOHYD 396 396 O-LINKED (GLCNAC. .) (PROBABLE).
FT CARBOHYD 453 453 O-LINKED (GLCNAC. .) (PROBABLE).
FT CARBOHYD 494 494 N-LINKED (GLCNAC. .) (PROBABLE).
FT ACT_SITE 434 434 CHARGE RELAY SYSTEM.
FT ACT_SITE 483 483 CHARGE RELAY SYSTEM.
FT ACT_SITE 578 578 CHARGE RELAY SYSTEM.
FT DISULFID 21 104 BY SIMILARITY.
FT DISULFID 47 77 BY SIMILARITY.
FT DISULFID 51 57 BY SIMILARITY.
FT DISULFID 111 194 BY SIMILARITY.
FT DISULFID 137 166 BY SIMILARITY.
FT DISULFID 141 147 BY SIMILARITY.
FT DISULFID 201 284 BY SIMILARITY.
FT DISULFID 227 256 BY SIMILARITY.
FT DISULFID 231 237 BY SIMILARITY.
FT DISULFID 292 375 BY SIMILARITY.
FT DISULFID 318 347 BY SIMILARITY.
FT DISULFID 322 328 BY SIMILARITY.
FT DISULFID 340 345 BY SIMILARITY.
FT DISULFID 383 503 BY SIMILARITY.
FT DISULFID 419 435 BY SIMILARITY.
FT DISULFID 517 584 BY SIMILARITY.
FT DISULFID 548 563 BY SIMILARITY.
FT DISULFID 574 602 BY SIMILARITY.
SQ SEQUENCE 638 AA; 71368 MW; CC27C93AB1086599 CRC64;

Alignment Scores:
Pred. No.: 1.89e-35 Length: 638
Score: 464.00 Matches: 95
Percent Similarity: 53.49% Conservative: 43
Best Local Similarity: 36.82% Mismatches: 102
Query Match: 27.82% Indels: 18
DB: 1 Gaps: 6

US-09-735-713a-1 (1-921) x KAL_MOUSE (1-638)
QY 154 ATTTTCAGTCGCATTTCTGGAGGAAGCCAAAGTGGAGAGGTTCTATCCCTGGCAGGTA 213
Db 387 IleAsnAlaArgIleValGlyGlyThrAsnAlaSerLeuGlyGluTrpProTrpGlnVal 406
QY 214 TCTCTGAACAACAAAGG-----CAGAAGCATATTTGTGGAGGAAGCATCTCTCACA 264
Db 407 SerLeuGlnValLysLeuValSerGlnThrHisLeuCysGlyGlySerIleIleGlyArg 426
QY 265 CAGTGGGTGATCACCGCGCTCAGTCGATTCGAAACAGAAACATGTGTCTACTTTGAAT 324
Db 427 GlnTrpValLeuThrAlaAlaHisCysPheAspGlyIleProTyrProAspValTrpArg 446

```

```
QY 325 GTTACTGCTGGAGATGACTTAAAGCCAGACAGACCCAGGAGCAAACTCTCACTATT 384
DB 447 lletyGlyGlyIleLeuSerLeuSerGluIleThrLysGluThrProSerSerArgile 466
QY 385 GAAGCTGTCATCATACATCCACATTCCTCCACCAAGAACCAAGGATATGATATGCC 444
DB 467 LysGluLeuIleIleHisGlnGluThrLysValSerGluGly--AsnTyrAspIleAla 485
QY 445 CTTTGAAGATGGCTGGAGCCCTTCAATTTGGCCACTTTGGGGCCCATATGCTTCCA 504
DB 486 LeuIleLysLeuGlnThrProLeuAsnTyrThrGluPheGlnLysProIleCysLeuPro 505
QY 505 GAGCTGCGGGAGCAATTTGAGCGCTGTTTATTGTACAACTGCAGGCTGGCGCGCTTA 564
DB 506 SerLysAlaAspThrAsnThrIleTyrThrAsnCysTyrPValThrGlyTyrThr 525
QY 565 ACTGAAGTGGCGCTCTCACAAGTCTTGAGGAGTGAATCGCTATTTGACCTGG 624
DB 526 LysGluGlnGlyGluThrGlnAsnIleLeuGlnLysAlaThrIleProLeuValProAsn 545
QY 625 GAAGAGTGTGTGGCAGCTCTGTAACTAAAGAGGCCCATCAGTGGGAAGACCTTCTT 684
DB 546 GluGluCysGlnLys-----LysTyrArgAspTyrValIleAsnLysGlnMetile 562
QY 685 TGCACAGCTTTTCTGATGAGGAGACGACATGTCAGGAGATTCAGGAGGTTCACTC 744
DB 563 CysAlaGlyTyrLysGluGlyThrAspAlaCysLysGlyAspSerGlyGlyProLeu 582
QY 745 ATGTCCCGATAGAAGAGGCGCTGGACTCGCTGTGTGACATTCCTGGGTTTGGCC 804
DB 583 ValCysLysHis--SerGlyArgTyrGlnLeuValGlyIleThrSerTyrGlyGluGly 601
QY 805 TGTGTCTCAGGCTGGGAGAAACAATGTGAGAAAGTCAAGGATCCCTCGGGATCTTC 864
DB 602 CysGlyArg-----LysAspGln-----ProGlyValTyr 611
QY 865 ACAGACATTAGTAAGTCTTCTGTCGATCCAGCAACATCCAACTGGTAAC 918
DB 612 ThrLysValSerGluTyrMetAspTyrIleLeuGluLysThrGlnSerSerAsp 629
RESULT 4
TMS3_HUMAN
ID TMS3_HUMAN STANDARD; PRT; 454 AA.
AC P57727;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transmembrane protease, serine 3 (EC 3.4.21.-) (Serine protease
DE TADG-12) (Tumor associated differentially-expressed gene-12 protein).
GN TMPSR3 OR TADG12 OR ECHOS1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN RP SEQUENCE FROM N.A. (ISOFORMS A AND TRUNCATED).
RC TISSUE-Ovarian carcinoma;
RX MEDLINE=20521358; PubMed=11069177;
RA Underwood L.J., Shigemasa K., Tanimoto H., Beard J.B., Schneider E.N.,
RA Wang Y., Parmley T.H., O'Brien T.J.;
RT "Ovarian tumor cells express a novel multi-domain cell surface serine
RT protease.";
RL Biochim. Biophys. Acta 1502:337-350(2000).
[2]
RN RP SEQUENCE FROM N.A. (ISOFORMS A; B/C AND D), AND VARIANT ILE-53.
RX MEDLINE=20578749; PubMed=11137999;
RA Scott H.S., Kudoh J., Wattenhofer M., Shibuya K., Berry A., Chraet R.,
RA Guipponi M., Wang J., Kawasaki K., Asakawa S., Minoshima S.,
RA Youns F., Mehdi S.O., Radhakrishna U., Papasavvas M.P., Gehrig C.,
RA Rossier C., Korostishevsky M., Gal A., Shimizu N., Bonne-Tamir B.,
RA Antonarakis S.E.;
RT "Insertion of beta-satellite repeats identifies a transmembrane
RT protease causing both congenital and childhood onset autosomal
```

```
RT recessive deafness.";
RL Nat. Genet. 27:59-63(2001).
[3]
RN VARIANTS DFBN8/DFNB10 CYS-251 AND LEU-404.
RX MEDLINE=21354482; PubMed=11462234;
RA Masoudi S., Antonarakis S.E., Schwede T., Ghorbel A.M., Gratri M.,
RA Papasavvas M.P., Drira M., Elgaled-Boullila A., Wattenhofer M.,
RA Rossier C., Scott H.S., Ayadi H., Guipponi M.;
RT "Novel missense mutations of TMPSR3 in two consanguineous Tunisian
RT families with non-syndromic autosomal recessive deafness.";
RL Hum. Mutat. 18:101-108(2001).
[4]
RN VARIANTS DFBN8/DFNB10 TRP-109; PHE-194 AND ARG-407, AND VARIANTS
RX ILE-53; SER-111 AND VAL-253.
RX PubMed=11424922;
RA Ben-Yosef T., Wattenhofer M., Riazuddin S., Ahmed Z.M., Scott H.S.,
RA Kudo J., Shibuya K., Antonarakis S.E., Bonne-famil B.,
RA Radhakrishna U., Naz S., Ahmed Z., Riazuddin S., Pandya A.,
RA Nance W.E., Wilcox E.R., Friedman T.B., Morell R.J.;
RT "Novel mutations of TMPSR3 in four DFBN8/B10 families segregating
RT congenital autosomal recessive deafness.";
RL J. Med. Genet. 38:396-400(2001).
[5]
RN VARIANT DFBN8/DFNB10 GLY-103, AND VARIANTS ASN-173 AND THR-426.
RX PubMed=11907649;
RA Wattenhofer M., Di Iorio V., Rabionet R., Dougherty L., Pampanos A.,
RA Schwede T., Montserrat-Sentis B., Arbones L., Illades T.,
RA Pasquaddibisceglie A., D'Amelio M., Alwan S., Rossier C., Dahl H.-H.M.,
RA Petersen M.B., Estivill X., Gasparini P., Scott H.S.,
RA Antonarakis S.E.;
RT "Mutations in the TMPSR3 gene are a rare cause of childhood
RT nonsyndromic deafness in Caucasian patients.";
RL J. Mol. Med. 80:124-131(2002).
CC - FUNCTION: Probable protease.
CC - SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC - ALTERNATIVE PRODUCTS: 4 ISOFORMS; A (SHOWN HERE), B/C, D AND
CC TRUNCATED/TADG-12V; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC - TISSUE SPECIFICITY: Expressed in many tissues including fetal
CC cochlea. The truncated isoform is found at increased levels in
CC some carcinomas.
CC - DISEASE: Defects in TMPSR3 are a cause of two forms of autosomal
CC neurosensory childhood-onset forms of deafness, DFBN8 and DFBN10.
CC - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC - SIMILARITY: CONTAINS 1 SRCR DOMAIN.
CC - SIMILARITY: CONTAINS 1 LDL-RECEPTOR CLASS A DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
[-----]
EMBL; AF201380; AAG37012.1; -
DR EMBL; AB038157; BAB20077.1; -
DR EMBL; AB038158; BAB20078.1; -
DR EMBL; AB038159; BAB20079.1; -
DR EMBL; AB038160; BAB20080.1; -
DR HSSP; P00763; 1DPO.
DR MEROPS; S01.079; -
DR GeneW; HGNC:11877; TMPSR33.
DR MIM; 605511; -
DR MIM; 601072; -
DR MIM; 605316; -
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR002172; LDL_receptL.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00057; ldl_recept_a; 1.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00192; LDLa; 1.
```

DR SMART; SM00202; SR; 1.  
DR SMART; SM00202; Tryp\_Spc; 1.  
DR PROSITE; PS01209; LDLRA\_1; 1.  
DR PROSITE; PS00068; LDLRA\_2; 1.  
DR PROSITE; PS00420; SRCR\_1; FALSE\_NEG.  
DR PROSITE; PS00287; SRCR\_2; 1.  
DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE; PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
KW Hydrolase; Serine protease; Transmembrane; Signal-anchor; Deafness;  
KW Alternative splicing; Disease mutation; Polymorphism.  
FT DOMAIN 1 48 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 49 69 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
(POTENTIAL).  
FT DOMAIN 70 454 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 72 108 LDL-RECEPTOR CLASS A.  
FT DOMAIN 109 205 SRCR.  
FT DOMAIN 217 454 SERINE PROTEASE.  
FT ACT\_SITE 257 257 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 304 304 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 401 401 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT SITE 216 217 CLEAVAGE (POTENTIAL).  
FT DISULFID 73 85 BY SIMILARITY.  
FT DISULFID 79 98 BY SIMILARITY.  
FT DISULFID 92 107 BY SIMILARITY.  
FT DISULFID 129 194 BY SIMILARITY.  
FT DISULFID 142 204 BY SIMILARITY.  
FT DISULFID 207 324 BY SIMILARITY.  
FT DISULFID 242 258 BY SIMILARITY.  
FT DISULFID 338 407 BY SIMILARITY.  
FT DISULFID 370 386 BY SIMILARITY.  
FT DISULFID 397 425 BY SIMILARITY.  
FT CARBOHYD 221 127 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARSPLIC 1 127 MISSING (IN ISOFORM B/C).  
FT VARSPLIC 318 454 MISSING (IN ISOFORM B/C).  
FT EMLOPCLPNSERNRDPDGKVCWTSGMGATEDGAGDASPVLN  
HAAPLISNKHNRDVGGLISPSMLCAGLYGGVDSQCG  
DSGGPLVCERRLWKLKVGATSGIGCAEVNKPVGYPRTVSF  
LWMIHQMERDLTK -> DTSGSLGSAALPLFQEDQLLI  
EAPL (IN ISOFORM D).  
FT DLYLPKSTIQVGLSLDNPAPSHLVEKIVYH -> EIVA  
PRERADRRKLLCNKPKTKMKGPRPSHS (IN  
TRUNCATED ISOFORM).  
FT V -> I.  
/FTID=VAR\_010781.  
D -> G (IN DFN88/DFNB10).  
/FTID=VAR\_013490.  
R -> W (IN DFN88/DFNB10).  
/FTID=VAR\_013491.  
G -> S.  
/FTID=VAR\_013492.  
D -> N.  
/FTID=VAR\_013493.  
C -> F (IN DFN88/DFNB10).  
/FTID=VAR\_013494.  
W -> C (IN DFN88/DFNB10).  
/FTID=VAR\_011678.  
I -> V (IN DBSNP:2839500).  
/FTID=VAR\_013101.  
P -> L (IN DFN88/DFNB10).  
/FTID=VAR\_011679.  
C -> R (IN DFN88/DFNB10).  
/FTID=VAR\_013495.  
A -> T.  
/FTID=VAR\_013496.  
LAFEPPLIVI -> FEVESOSSSL (IN REF. 1).  
A -> T (IN REF. 1).  
MISSING (IN REF. 1).  
FT ICNHRDVGGLISPSMLCAGLYTGGVD -> DLQPQGRVRW  
HHPLHLALRGLDPDGRWN (IN REF. 1).  
E -> D (IN REF. 1).  
454 AA; 49404 MW; 57ECC3678F7D6AFF CRC64;  
SQ SEQUENCE

Alignment Scores: 8.84e-35 Length: 454  
Pred. No.: 456.50 Matches: 99  
Score: 51.78% Conservative: 61  
Percent Similarity: 32.04% Mismatches: 114  
Best Local Similarity: 27.37% Indels: 35  
Query Match: 1  
DB: 10  
US-09-735-713a-1 (1-921) x TMS3\_HUMAN (1-454)  
QY 1 ATGAGTCTCAAAATGCTTATAAGCAGAACAGCTGATTTACTACTAGGAATAGTCTTT 60  
Db 170 ValSerIleAspHisLeuLeuProAspLysValThrAlaLeuHisHisSerValTyr 189  
QY 61 TTTGAACRAGGT--AAATCTGCARCTCTTCGCTCCCAAGCTCCCACTTTGGGCAG 117  
Db 190 ValArgGluGlyCysAlaSerGlyHisValValThrLeuGlnCysThrAlaCysGlyHis 209  
QY 118 AGTCTGGTTAAGGTACAGCCTTGGAAATATTATTAACATTTTTCAGTCGATTCCTGGAGGA 177  
Db 210 ArgArg-----GlyTyrSerSerArgIleValGlyGly 220  
QY 178 AGCCAACTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAACAAAGGAGAGCAT 237  
Db 221 AsnMetSerLeuLeuSerGlnTrpTrpGlnAlaSerLeuGlnPheGlnGlyTyrHis 240  
QY 238 ATTTGTGGAGAAAGCATCGTCTCACCACAGTGTATCAGCGGCTCACTGCTCATTCGCCA 297  
Db 241 LeuCysGlyGlySerValIleThrProLeuTrpIleIleThrAlaAlaHisCysValTyr 260  
QY 298 AACAGAAACATTTGCTCTACTTTGAATGTTACTGCTGAGAGATGACTTAAGCCAGACA 357  
Db 261 AspLeuTyrLeuProLysSerTrpThrIleGlnValGlyLeuValSerLeuLeuAsp-- 279  
QY 358 GACCCAGGAGAGCAAACTCTCACTATTGAACACTGTATCATCATCATCATTCCTCCACC 417  
Db 280 AsnProAlaProSerHisLeu--ValGluLysIleValTyrHisSerLysTyrLysPro 298  
QY 418 AAGAAACCAATGACTATGATATTGCCTTTTGAAGATGGCTGGAGCTTCCAATTTGGC 477  
Db 299 LysArg--LeuGlyAsnAspIleAlaLeuMetLysLeuAlaGlyProLeuThrPheAsn 317  
QY 478 CACTTTGTGGGCCCATATGATTTCTTCAGAGCTCGGGAGCAATTTGAGCTGCTTTTATT 537  
Db 318 GluMetIleGlnProValCysLeuProAsnSerGluGluAsnPheProAspGlyLysVal 337  
QY 538 TGTACAACTGCAGGCTGGGGCCCTTAACCTGAAGGT---GGCGTCTCTCACAGTCTTG 594  
Db 338 CysTrpThrSerGlyTyrGlyAlaThrGluAspGlyAlaGlyAspAlaSerProValLeu 357  
QY 595 CAGGAAGTGAATCTGCCTATTTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTTAACACTA 654  
Db 358 AsnHisAlaAlaValProLeuIleSerAsnLysIleCys-----Asn 371  
QY 655 AAGAGCCCATCATGCTGGG-----AAGACTTTCTTTGACAGAGTTTCTCGAT 702  
Db 372 HisArgAspValTyrGlyGlyIleIleSerProSerMetLeuCysAlaGlyTyrLeuThr 391  
QY 703 GGAGGAGAGAGCCATGTCAGGAGATTCAGGAGGTTCATCATGTCGCGGAATAGAA 762  
Db 392 GlyGlyValAspSerGlnGlyAspSerGlyGlyProLeuValCysGlnGluArg 411  
QY 763 GGGGCTGGACTCTGCTGCTGTGACTTCTCTGGGTTTGGCTGTGGTCTGAGGCTGAGA 822  
Db 412 ---LeuTrpLysLeuValGlyAlaThrSerPheGlyIleGlyCys----- 425  
QY 823 AACATGTGAGGAAAGTGAATCAAGGATCCCTGGGATCTTCACAGACATTAGTAAAGTG 882  
Db 426 -----AlaGluValAsnLysProGlyValTyrThrArgValThrSerPhe 440  
QY 883 CTTTCTGGATCCACCAACACATCCAA 909  
Db 441 LeuAspTrpIleHisGluGlnMetGlu 449





CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91129236; PubMed=1993180;  
 RA Beaubien G., Rosinski-Chupin I., Mattei M.-G., Mbikay M., Chretien M.,  
 RA Seidah N.G.;  
 RT "Gene structure and chromosomal localization of plasma kallikrein.";  
 RL Biochemistry 30:1628-1635(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=90091743; PubMed=2598771;  
 RA Seidah N.G., Ladenheim R., Mbikay M., Hamelin J., Lutfalla G.,  
 RA Rougeon F., Lazure C., Chretien M.;  
 RT "The cDNA structure of rat plasma kallikrein.";  
 RL DNA 8:563-574(1989).  
 CC -|- FUNCTION: THE ENZYME CLEAVES LYS-ARG AND ARG-SER BONDS. IT  
 CC ACTIVATES, IN A RECIPROCAL REACTION, FACTOR XII AFTER ITS BINDING  
 CC TO A NEGATIVELY CHARGED SURFACE. IT ALSO RELEASES BRADYKININ FROM  
 CC HMW KININOGEN AND MAY ALSO PLAY A ROLE IN THE RENIN-ANGIOTENSIN  
 CC SYSTEM BY CONVERTING PRORENIN INTO RENIN.  
 CC -|- CATALYTIC ACTIVITY: Cleaves selectively Arg-I-Xaa and Lys-I-Xaa  
 CC bonds, including Lys-I-Arg and Arg-I-Ser bonds in (human)  
 CC kininogen to release bradykinin.  
 CC -|- SUBUNIT: THE ZYMOGEN IS ACTIVATED BY FACTOR XIIA, WHICH CLEAVES  
 CC THE MOLECULE INTO A LIGHT CHAIN, WHICH CONTAINS THE ACTIVE SITE,  
 CC AND A HEAVY CHAIN, WHICH ASSOCIATES WITH HMW KININOGEN. THESE  
 CC CHAINS ARE LINKED BY ONE OR MORE DISULFIDE BONDS.  
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMA KALLIKREIN  
 CC SUBFAMILY.  
 CC -|- SIMILARITY: CONTAINS 4 APPLE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M62357; AAA74563.1; -;  
 CC EMBL; M62358; AAA74563.1; JOINED.  
 CC EMBL; M62346; AAA74563.1; JOINED.  
 CC EMBL; M62347; AAA74563.1; JOINED.  
 CC EMBL; M62349; AAA74563.1; JOINED.  
 CC EMBL; M62350; AAA74563.1; JOINED.  
 CC EMBL; M62351; AAA74563.1; JOINED.  
 CC EMBL; M62352; AAA74563.1; JOINED.  
 CC EMBL; M62353; AAA74563.1; JOINED.  
 CC EMBL; M62354; AAA74563.1; JOINED.  
 CC EMBL; M62355; AAA74563.1; JOINED.  
 CC EMBL; M62356; AAA74563.1; JOINED.  
 CC EMBL; M30282; AAA41463.1; -;  
 CC EMBL; M58590; AAA42069.1; -;  
 CC PIR; A39180; KQRTPL.  
 CC HSSP; P00750; IRTF.  
 CC MEROPS; S01.212; -;  
 CC InterPro; IPR000177; Apple.  
 CC InterPro; IPR001314; Chymotrypsin.  
 CC InterPro; IPR003014; PAN.  
 CC InterPro; IPR001254; Ser\_protease\_Try.  
 CC Pfam; PF00024; PAN; 4.  
 CC Pfam; PF00089; trypsin; 1..  
 CC PRINTS; PR00005; APPLIEDOMAIN.  
 CC PRINTS; PR00722; CHYMOTRYPSIN.  
 CC SMART; SM00223; APPLE; 4.  
 CC SMART; SM00020; Tryp\_SPC; 1.  
 CC PROSITE; PS00495; APPLE; 4.  
 CC PROSITE; PS02040; TRYPSIN\_DOM; 1.  
 CC PROSITE; PS00134; TRYPSIN\_HIS; 1.  
 CC PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase: Serine protease; Glycoprotein; Plasma; Zymogen; Signal;  
 KW Fibrinolysis; Blood coagulation; Inflammatory response; Liver;

KW Repeat. 1 19  
 FT SIGNAL 20 390 PLASMA KALLIKREIN HEAVY CHAIN.  
 FT CHAIN 391 638 PLASMA KALLIKREIN LIGHT CHAIN.  
 FT CHAIN 391 638  
 FT DOMAIN 20 105 APPLE 1.  
 FT DOMAIN 110 195 APPLE 2.  
 FT DOMAIN 200 285 APPLE 3.  
 FT DOMAIN 291 376 APPLE 4.  
 FT DOMAIN 389 621 SERINE PROTEASE.  
 FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (PROBABLE).  
 FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (PROBABLE).  
 FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (PROBABLE).  
 FT CARBOHYD 453 453 N-LINKED (GLCNAC. . .) (PROBABLE).  
 FT CARBOHYD 494 494 N-LINKED (GLCNAC. . .) (PROBABLE).  
 FT ACT\_SITE 434 434 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 483 483 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 578 578 CHARGE RELAY SYSTEM.  
 FT DISULFID 21 104 BY SIMILARITY.  
 FT DISULFID 47 77 BY SIMILARITY.  
 FT DISULFID 51 57 BY SIMILARITY.  
 FT DISULFID 111 194 BY SIMILARITY.  
 FT DISULFID 137 166 BY SIMILARITY.  
 FT DISULFID 141 147 BY SIMILARITY.  
 FT DISULFID 201 284 BY SIMILARITY.  
 FT DISULFID 227 256 BY SIMILARITY.  
 FT DISULFID 231 237 BY SIMILARITY.  
 FT DISULFID 292 375 BY SIMILARITY.  
 FT DISULFID 318 347 BY SIMILARITY.  
 FT DISULFID 322 328 BY SIMILARITY.  
 FT DISULFID 340 345 BY SIMILARITY.  
 FT DISULFID 383 503 BY SIMILARITY.  
 FT DISULFID 419 435 BY SIMILARITY.  
 FT DISULFID 517 584 BY SIMILARITY.  
 FT DISULFID 548 563 BY SIMILARITY.  
 FT DISULFID 574 602 BY SIMILARITY.  
 SQ SEQUENCE 638 AA; 71273 MW; 454BEB27E8CA8F88 CRC64;  
 Alignment Scores:  
 Pred. No.: 9.14e-34 Length: 638  
 Score: 446.00 Matches: 91  
 Percent Similarity: 53.12% Conservatve: 45  
 Best Local Similarity: 35.55% Mismatches: 102  
 Query Match: 26.74% Indels: 18  
 DB: 1 Gaps: 5  
 US-09-735-713A-1 (1-921) x KAL\_RAT (1-638)  
 Qy 154 ATTTTCAGTCGATCTTGGAGGAGCCAGTGGAGAGGGTTCCTATCTCTGCGGAGGTA 213  
 Db 387 IleAsnAlaArgIleValGlyGlyThrAsnSerSerLeuGlyGluTrpProTrpGlnVal 406  
 Qy 214 TCTCTGAACAAGG-----CAGAAGCATATTTCTGGAGGAGCATGCTCTCACCA 264  
 Db 407 SerLeuGlnValLysLeuValSerGlnAsnHisMetCysGlyGlySerIleIleGlyArg 426  
 Qy 265 CAGTGGGTGATCACGGCGCTCAGTCATTCGCAACAGAAACATTTGTCTACTTTGAAT 324  
 Db 427 GlnTrpIleLeuThrAlaAlaHisCysPheAspGlyIleProTrpProaspValTrpArg 446  
 Qy 325 GTTACTGCTGGAGAGTATGACTTAAGCCAGCAGACAGCCAGGAGAGCAACTCTCACTATT 384  
 Db 447 IleTyrglyGlyIleLeuAsnLeuSerGluIleThrAsnLysThrProPheSerSerIle 466  
 Qy 385 GAAACTGTCAATACATCCACATTCCTCCACCAAGAAACCAATGATATGATATTCGCC 444  
 Db 467 LysGluLeuIleIleHisGlnLysTyrlLysMetSerGlyGly---SerTyrlLysIleAla 485  
 Qy 445 CTTTGAAGATGGCTGGAGCCCTTCCAATTTGGCCACTTTTGGGGCCCATATGCTTCCA 504  
 Db 486 LeuIleLysLeuGlnThrProLeuAsnTyThrGluPheGlnLysProIleCysLeuPro 505  
 Qy 505 GAGTGGCGGGAGCAATTTGAGCTGGTGTATTATTCTACAACTGACAGCTGGGGCGCGTTA 564  
 Db 505 GAGTGGCGGGAGCAATTTGAGCTGGTGTATTATTCTACAACTGACAGCTGGGGCGCGTTA 564



FT ACT\_SITE 987 987 CHARGE RELAY SYSTEM (BY SIMILARITY).  
 FT LIPID 2 2 MYRISTATE (POTENTIAL).  
 FT DISULFID 199 212 BY SIMILARITY.  
 FT DISULFID 206 225 BY SIMILARITY.  
 FT DISULFID 219 236 BY SIMILARITY.  
 FT DISULFID 659 671 BY SIMILARITY.  
 FT DISULFID 666 684 BY SIMILARITY.  
 FT DISULFID 678 693 BY SIMILARITY.  
 FT DISULFID 788 812 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 826 842 BY SIMILARITY.  
 FT DISULFID 926 993 BY SIMILARITY.  
 FT DISULFID 957 972 BY SIMILARITY.  
 FT DISULFID 983 1011 BY SIMILARITY.  
 FT CARBOHYD 116 116 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 147 147 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 170 170 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 194 194 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 233 233 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 263 263 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 264 264 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 404 404 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 456 456 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 486 486 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 519 519 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 550 550 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 646 646 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 698 698 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 722 722 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 741 741 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 762 762 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 864 864 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 903 903 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 965 965 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT VARSPLIC 166 192 MISSING (IN SHORT ISOFORM).  
 FT CONFLICT 808 808 R -> Y (IN REF. 3).  
 SQ SEQUENCE 1035 AA; 114887 MW; E207970B08296E13 CRC64;  
  
 Alignment Scores:  
 Pred. No.: 2.15e-33 Length: 1035  
 Score: 442.50 Matches: 96  
 Percent Similarity: 54.55% Conservative: 54  
 Best Local Similarity: 34.91% Mismatches: 90  
 Query Match: 26.53% Indels: 35  
 DB: 1 Gaps: 9  
  
 US-09-735-713a-1 (1-921) x ENT\_K\_BOVIN (1-1035)  
 QY 106 AGTTGGGCGAGCTGGTT-----AAGTACAGCCTTGAATATTATTTAAACATTTTC 159  
 Db 787 SerCysGlyLysLysLeuValThrGlnGluValSerPro----- 799  
 QY 160 AGTCGATCTTGGAGGAGCAAGTGGAGAGGGTTCCTATCCCTGGCAGGTACTCTGT 219  
 Db 800 ---LysIleValGlySerAspSerArgGluGlyAlaIrrProTrpValValAlaLeu 818  
 QY 220 AAACAAAGCAGAGCATATTGTGGAGGAGCATCGTCTCACCACAGTGGGTGATCAGC 279  
 Db 819 TyrPheAspGlnGlnValCysGlyAlaSerLeuValSerArgAspTrpLeuValSer 838  
 QY 280 GCGGCTCACTGCATTCGAAACAGAAACATTT---GTGCTACTTTGAATGTTACTGCTGGA 336  
 Db 839 AlaAlaHisCysValTyrGlyArgAsnMetGluProSerLysTrpLysAlaValLeuGly 858  
 QY 337 GAGTATGACTTAAGCCAG---ACAGACCCAGGAGCAAACTCTCATTATGAACACTGC 393  
 Db 859 LeuHisMetAlaSerAsnLeuThrSerProGlnIleGluThrArgLeuIleAspGlnIle 878  
 QY 394 ATCATACATCACATTTCTCCACCAAGAAACCAATGGACTATGATATTCGCTTTTGAAG 453  
 Db 879 VallLeuAsnProHisTyrAsn---LysArgArgLysAsnAspIleAlaMetMetHis 897  
 QY 454 ATGCTGGAGCCTTCCCAATTTGGCCACTTTGTGGGGCCCATATGTTCTCCAGAGCTGGG 513  
 Db 454 ATGCTGGAGCCTTCCCAATTTGGCCACTTTGTGGGGCCCATATGTTCTCCAGAGCTGGG 513

Db 898 LeuGluMetLysValAsnTyrThrAspTyrIleGlnProIleCysLeuProGluGluAsn 917  
 QY 514 GAGCAATTTGAGGCTGGTTTATTTGTACAACTGCAGCTGGGGCGCTTAACTGAAGT 573  
 Db 918 GlnValPheProGlyArgIleCysSerIleAlaGlyTrpGlyAlaLeuIleTyrGln 937  
 QY 574 GCGCTCCCTCTCAGGAAGTGAATCTGCCTATTTTACCTGGGAAGAGTGT 633  
 Db 938 GlySerThrAlaAspValLeuGlnGluAlaAspValProLeuLeuSerAsnGlnLysCys 957  
 QY 634 GTGGCAGCTCTG-----TTAACACTAAAGAGCCCATCAGTGGGAAGACCTTT 681  
 Db 958 GlnGlnGlnMetProGluTyrAsnIleThr-----GluAsnMet 970  
 QY 682 CTTTCACAGCTTTTCTGATGGAGGAGAGCAGCATCTCAGGAGATTCAGGAGTTCA 741  
 Db 971 ValCysAlaGlyTyrGlnAlaGlyValAspSerCysGlnGlyValSerGlyGlyPro 990  
 QY 742 CTCATGTCGCCGAATAAGAAAGGCGCTGAGTCTGCGTGTGTGACTTCCTGGGTTTG 801  
 Db 991 LeuMetCysGln---GluAsnAsnArgTrpLeuLeuAlaGlyValThrSerPheGlyTyr 1009  
 QY 802 GCCTGTGCTCGAGGCTGGAGAAACAATCTGAGGAAAGTGTCAAGGATCCCTGGATC 861  
 Db 1010 GlnCysAlaLeuProAsnArg-----ProGlyVal 1019  
 QY 862 TTCACAGACATTAGTAAAGTCTTTCCTGGATCCACGAACACATC 906  
 Db 1020 TyrAlaArgValProArgPheThrGluTrpIleGlnSerPheLeu 1034  
 RESULT 9  
 PSS8\_HUMAN STANDARD; PRT; 343 AA.  
 ID O16651; O9UCA3;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Proctasin precursor (EC 3.4.21.-).  
 GN PRSS8.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Prostate;  
 RX MEDLINE=95285644; PubMed=7768952;  
 RT Yu J.X., Chao L., Chao J.;  
 RT "Molecular cloning, tissue-specific expression, and cellular  
 RT localization of human proctasin mRNA.";  
 RL J. Biol. Chem. 270:13483-13489(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX Strausberg R.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 45-64.  
 RC TISSUE=Semen;  
 RX MEDLINE=94308140; PubMed=8034638;  
 RA Yu J.X., Chao L., Chao J.;  
 RT "Proctasin is a novel human serine proteinase from seminal fluid.  
 RT Purification, tissue distribution, and localization in prostate  
 RT gland.";  
 RL J. Biol. Chem. 269:18843-18848(1994).  
 CC -!- FUNCTION: POSSESSES A TRYPSIN-LIKE CLEAVAGE SPECIFICITY.  
 CC -!- SUBUNIT: HETERODIMER OF TWO CHAINS, LIGHT AND HEAVY, HELD BY A  
 CC DISULFIDE BOND.  
 CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. SECRETED AFTER CLEAVAGE OF  
 CC ITS C-TERMINUS.  
 CC -!- TISSUE SPECIFICITY: FOUND IN PROSTATE, LIVER, SALIVARY GLAND,  
 CC KIDNEY, LUNG, PANCREAS, COLON, BRONCHUS AND RENAL PROXIMAL TUBULAR  
 CC CELLS. IN THE PROSTATE GLAND IT MAY BE SYNTHESIZED IN EPITHELIAL

CC CELLS, SECRETED INTO THE DUCTS, AND EXCRETED INTO THE SEMINAL  
 CC FLUID.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; L41351; AAC41759.1; -;  
 DR EMBL; U33446; AAB19071.1; -;  
 DR EMBL; BC001462; AAH01462.1; -;  
 DR HSSP; P00763; IDPO.  
 DR MEROPS; S01.159; -;  
 DR Genew; HGNC:9491; PRSS8.  
 DR MIM; 600823; -;  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR001254; Ser.protease\_Try.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; TRYP\_SPC; 1.  
 DR PROSITE; PS0240; TRYPsin\_DOM; 1.  
 DR PROSITE; PS00134; TRYPsin\_HIS; 1.  
 DR PROSITE; PS00135; TRYPsin\_SER; 1.  
 DR Hydrolase; Serine protease; Zymogen; Signal; Glycoprotein;  
 KW Transmembrane.  
 FT SIGNAL 1 29 POTENTIAL.  
 FT PROPEP 30 32 ACTIVATION PEPTIDE.  
 FT CHAIN 33 44 PROSTATIN LIGHT CHAIN.  
 FT CHAIN 33 44 PROSTATIN HEAVY CHAIN.  
 FT PROPEP 323 343  
 FT PROPEP 323 343  
 FT TRANSXEM 320 340  
 FT DOMAIN 45 286  
 FT DISULFID 37 154 SERINE PROTEASE.  
 FT DISULFID 70 86 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 168 244 BY SIMILARITY.  
 FT DISULFID 201 223 BY SIMILARITY.  
 FT DISULFID 234 262 BY SIMILARITY.  
 FT ACT\_SITE 85 95 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 134 134 CHARGE RELAY SYSTEM.  
 FT ACT\_SITE 238 238 CHARGE RELAY SYSTEM.  
 FT CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 343 AA; 36431 MW; 98DD6447F5A8C1B2 CRC64;

## Alignment Scores:

Pred. No.: 2.35e-33 Length: 343  
 Score: 441.00 Matches: 100  
 Percent Similarity: 49.16% Conservative: 46  
 Best Local Similarity: 33.67% Mismatches: 117  
 Query Match: 26.44% Indels: 34  
 DB: 1 Gaps: 10

US-09-735-713a-1 (1-921) x PSS8\_HUMAN (1-343)

QY 34 CTGATTTTACTAGTAAGTCTTTTGAACRAGGTAAATCTGCARCTCTTTGCTC 93  
 Db 17 ILeuLeuTyrLeuGlyLeuLeuArgSerGlyThrGlyAlaGluGlyAla----- 33  
 QY 94 CCAAAAGCTCCCAAGTTGTGGCGAGAGCTCTGGTTAAGTACAGCTTGGAAATATTTTAAAC 153  
 Db 34 ---GluAlaPro---CysGly-----ValalaProGln----- 42  
 QY 154 ATTTTCAGTCGCAATTTGAGGAAGCCCAAGTGGAGAGGGTTCATCCCTGGCAGGTA 213  
 Db 43 -----AlaargilethrGlyGlySerAlaValalaGlyGlntrpGlnVal 60  
 QY 214 TCTCTGAACAAGCGAGAGCATATTTGGAGGAGGAGCATCTCTCACCACAGTGGTG 273  
 Db 61 SerilethrTyrGluGlyValHisValCysGlyGlySerLeuValSerGluGlntrpVal 80

QY 274 ATCAGCGCGCTCACTGCATTCGCAACAGAAACATTTGTCTACTTTGAATGTACTGCT 333  
 Db 81 LeuSerAlaAlaHisCysPheProSerGluHisHisLysGluAlaTyrGluValLysLeu 100  
 QY 334 GGAGAGTATGACTTAAGCCAGAGACACCAGAGAGCAAACTCTCACTATTGAACTGTC 393  
 Db 101 GlyAlaHisGlnLeuAspSerTyrSerGluAspAlaLysValSerThrLeuLysAspIle 120  
 QY 394 ATCATATCCACATTTCTCCCAAGAAACCAATGAGCATATGATATTCCTCTTTGAAG 453  
 Db 121 IleProHisProSerTyr---LeuGlnGluGlySerGlnGlyAspIleAlaLeuGln 139  
 QY 454 ATGGCTGGAGCCCTCCCAATTTGGCCACTTTGGCGCCATATGTCTCCAGAGTCGCG 513  
 Db 140 LeuSerArgProIleThrPheSerArgTyrIleArgProIleCysLeuProAlaAlaAsn 159  
 QY 514 GAGCAATTTGAGGCTGGTTTATTTGTCACACTGCAGGCTGGCGCGCTTAAGTGAAGT 573  
 Db 160 AlaSerPheProAsnGlyLeuHisCysThrValThrGlyTyrGlyHisValAlaProSer 179  
 QY 574 GCGTCTCTC-----TCACAAGTCTTGCAGGAAGTGAATCTGCCTATTTTGACCTGGAA 627  
 Db 180 ValSerLeuLeuThrProLysProLeuGlnGlnLeuGluValProLeuLeuSerArgGlu 199  
 QY 628 GAGTGTGTGGCAGCTCTGTAACTAAAGAGGCC-----ATCAGTGGGAAG 675  
 Db 200 ThrCysAsnCysLeuTyrAsnIleAspAlaLysProGluGluProHisPheValGlnGlu 219  
 QY 676 ACCTTTCTTTGACAGGTTTCTGTGATGAGGAGAGAGCATGTCTGATCCAGGAGATTCAGGA 735  
 Db 220 AspMetValCysAlaGlyTyrValGluGlyGlyLysAspAlaCysGlnGlyAspSerGly 239  
 QY 736 GGTTCACATCTGCGCGAATAAGAAAGGCGCTGCAGCTCTGCTGCTGCTGCTGCTGCTG 795  
 Db 240 GlyProLeuSerCys---ProValGluGlyLeuTyrTyrLeuThrGlyIleValSerTrp 258  
 QY 796 GGTGTGGCTGTGCTGAGGCTGGAGAAACAATGTGAGGAAAGTATCAAGGATCCCT 855  
 Db 259 GlyAspAlaCysGlyAla-----ArgAsnArg-----Pro 268  
 QY 856 GGAATCTTCACAGACATTAAGTAAAGTCTTCTGATCCAGCAACACATC 906  
 Db 269 GlyValTyrThrLeuAlaSerTyrAlaSerTrpIleGlnSerLysVal 285

RESULT 10  
 ENTK\_PIG STANDARD; PRT; 1034 AA.  
 ID ENTK\_PIG  
 AC P98074;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Enteroproteptidase precursor (EC 3.4.21.9) (Enterokinase).  
 GN PRSS7 OR ENTK.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Cetartiodactyla; Sulina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC TISSUE=Duodenal mucosa;  
 RX MEDLINE=94327548; PubMed=8051081;  
 RA Matsushima M., Ichinose M., Yahagi N., Kakei N., Tsukada S.,  
 RA Miki K., Kurokawa K., Tashiro K., Shiohara K., Shinomiya K.,  
 RA Uneyama H., Inoue H., Takahashi T., Takahashi K.;  
 RT "Structural characterization of porcine enteroproteptidase";  
 RL J. Biol. Chem. 269:19976-19982(1994).  
 CC -1- FUNCTION: RESPONSIBLE FOR INITIATING  
 CC PROTEOLYTIC PROENZYME (TRYPSIN, CHYMOTRYPSIN AND CARBOXYPEPTIDASE  
 CC A). IT CATALYZES THE CONVERSION OF TRYPSINOGEN TO TRYPSIN WHICH IN  
 CC TURN ACTIVATES OTHER PROENZYME INCLUDING CHYMOTRYPSINOGEN,  
 CC CARBOXYPEPTIDASES, AND PROELASTASES.  
 CC -1- CATALYTIC ACTIVITY: Selective cleavage of 6-Lys-|-Ile-7 bond in  
 CC trypsinogen.



```

Db 845 ArgAsnLeuGluProSerLysTrpLysAlaIleLeuGlyLeuHisMetThrSer----- 862
Qy 340 TATGACTTAAGCCAGACAGACAGGAGCAAACTCTCACTATTGTAAGAACTGTCATCATA 399
Db 863 -----AsnLeuThrSerProGlnIleValThrArgLeuIleAspGluIleValle 879
Qy 400 CATCACATTTCTCCACCAAGAACCAATGAGCTATGATATGCCCCCTTTGAAGATGGCT 459
Db 880 AsnProHisTyraAsnArgArg---LysAspSerAspIleAlaMetMetHisLeuGlu 898
Qy 460 GGACCTTCATTTGGCCACTTGTGGGCCCATATGCTTCCAGAGCTGCGGAGACAA 519
Db 899 PheLysValAsnTyThrAspTyrIleGlnProIleCysLeuProGluGluAsnGlnVal 918
Qy 520 TTTGAGGCTGTTTATTGTAACACTGCAGGCTGGGCGCTTAACCTGAAGTGGCGTC 579
Db 919 PheProGlyArgIleCysSerIleAlaGlyTrpGlyValIleTyThrGlnGlySer 938
Qy 580 CTCTCACAGTCTTGCAGGAAGTGAATCTGCTATTTGACCTGGGAGAGTGTGTGGCA 639
Db 939 ProAlaAspIleLeuGlnGluAlaAspValProLeuLeuSerAsnGluLysCysGlnGln 958
Qy 640 GCTCTG-----TTAACACTAAAGAGGCCCATCAGTGGGAGACCTTCTTTCG 687
Db 959 GlnMetProGluTyraAsnIleThr-----GluAsnMetMetCys 971
Qy 688 ACAGTTTCTGATGAGGAGAGACGATGTCAGGAGATTCAGGAGGTTTCATCTCATG 747
Db 972 AlaGlyTyraGluGluGlyGlyIleAspSerCysGlnGlyAspSerGlyGlyProLeuMet 991
Qy 748 TGCCGGATTAAGAAAGGGCTGACTCTGCTGGTGTGACTTCTTGGGGTGGGGTGT 807
Db 992 CysLeu---GluAsnAsnArgTrpLeuLeuAlaGlyValThrSerPheGlyTyraGlnCys 1010
Qy 808 GGTGAGGCTGGAGAAACAATGTGAGGAAAGTATCAAGATCCCTCGGATCTTCACA 867
Db 1011 AlaLeuProAsnArg-----ProGlyValTyraAla 1020
Qy 868 GACATTAGTAAGTGTCTTCTGATCCAGCAACATC 906
Db 1021 ArgValProLysPheThrGluTrpIleGlnSerPheLeu 1033

```

## RESULT 11

```

HEPS_MOUSE STANDARD; PRT; 416 AA.
ID HEP5_MOUSE
AC 035453;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serine protease hepsin (EC 3.4.21.-).
GN HPN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=98058912; PubMed=9395459;
RA Vu T.-K.H., Liu R.W., Haakma C., Tomasek J.J., Howard E.W.;
RT Identification and cloning of the membrane-associated serine
RT protease, hepsin, from mouse preimplantation embryos.
RL J. Biol. Chem. 272:31315-31320(1997).
CC -!- FUNCTION: PLAYS AN ESSENTIAL ROLE IN CELL GROWTH AND MAINTENANCE
CC OF CELL MORPHOLOGY.
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

```

DR EMBL; AF030065; AAB84221.1;
DR HSP; P00763; IDPO.
DR MEROPS; S01.224;
DR MGD; MGI:1196620; Hpn.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR01254; Ser_protease_Try.
DR InterPro; IPR001190; Sror_receptor.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00202; SR; 1.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS0240; TRYP_SIN_DOM; 1.
DR PROSITE; PS00134; TRYP_SIN_HIS; 1.
DR PROSITE; PS00135; TRYP_SIN_SER; 1.
DR KEGG; K01161; Serine protease; Transmembrane; Signal-anchor.
FT CHAIN 1 161 CATALYTIC CHAIN (POTENTIAL).
FT CHAIN 162 416 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1 16 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT TRANSMEM 17 43 (POTENTIAL).
FT DOMAIN 44 416 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 162 416 SERINE PROTEASE.
FT ACT_SITE 202 202 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 256 256 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 352 352 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 152 276 INTERCHAIN (BY SIMILARITY).
FT DISULFID 187 203 BY SIMILARITY.
FT DISULFID 321 337 BY SIMILARITY.
FT DISULFID 348 380 BY SIMILARITY.
FT CARBOHYD 111 111 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 416 AA; 44739 MW; 432194FF4004F848 CRC64;

```

## Alignment Scores:

```

Pred. No.: 3,04e-33 Length: 416
Score: 440.00 Matches: 99
Percent Similarity: 49.47% Conservative: 41
Best Local Similarity: 34.98% Mismatches: 99
Query Match: 26.38% Indels: 44
DB: 1 Gaps: 10

```

US-09-735-713A-1 (1-921) x HEP5\_MOUSE (1-416)

```

Qy 109 TGTGGGCGAGAGTCTGCTTAAGGTACAGCCTTGGAATATTTTAACTTTTCAGTCGCAIT 168
Db 152 CysGlyArgArgLysLeuProVal-----AspArgile 162
Qy 169 CTGGAGGAGCAAGTGGAGAGGTTCTCTATCCCTGGCAGGTATCTCTGAAACAAAGG 228
Db 163 ValGlyGlyClnAspSerSerLeuGlyArgTrpProTrpGlnValSerLeuArgTyraSp 182
Qy 229 CAGAACATATTTGTGGAGAGCATCTCTACACAGCTGGGTGTATCAGCGGCGTCA 288
Db 183 GlyThrHisLeuCysGlyGlySerLeuLeuSerGlyAspTrpValLeuThrAlaHis 202
Qy 289 TGCATTGCAACACAGAAC---ATTGCTCTACTTCTTGAATGTTACTCTGGAGAGTATGAC 345
Db 203 CysPheProGluArgAsnArgValLeuSerArgTrpArgValPheAlaAlyA----- 220
Qy 346 TTAAGCCAGACAGACCCAGGAGAGCAAACTCTCACTATTGAAACTGTCTATCATCATCA 405
Db 221 ValAlaArgThrSerProHisAlaValGlnLeuGlyValClnAlaValIleTyHisGly 240
Qy 406 CATTTT-----TCCACCAAGAACCAATGGACTATGATATTCCTTCCTTTT 450
Db 241 GlyTyraLeuProPheArgAspProThrIleAspGluAsnSerAsnAspIleAlaLeuVal 260
Qy 451 AAGATGGCTGGAGCCTTCCAAATTTGGCCACTTTGTGGGCCCATATGCTCTCCAGAGCTG 510
Db 261 HisLeuSerSerLeuProLeuThrGluTyraIleGlnProValCysLeuProAlaAla 280

```

```
QY 511 CGGAGCAATTGAGCGCTGTTTATTGTACAACTGCAGGCTGGCGCGCTTAACCTGAA 570
DB 281 GlyGlnAlaLeuValAspGlyLysValCysThrValThrGlyTyrpGlyAsnThrClnPhe 300
QY 571 GGTGGCTCTCTCACAAGTCTTCAGGAAGTGAATCGCCTATTGTCACCTGGGAAGAG 630
DB 301 TyrGlyGlnGlnAlaMetValLeuGlnGlnAlaArgValProIleIleSerAsnGluVal 320
QY 631 TGTGTGGCAGCTCTGTTAAACACTAAAGAGGCC-----ATCAGTGGG 672
DB 321 Cys-----AsnSerProAspPheTyrGlyAsnGlnIleLysPro 333
QY 673 AAGACCTTTCTTTCACAGGTTTCTGATGAGGAGGAGCGATGTCAGGGAGATTCAC 732
DB 334 LysMetPhe---CysAlaGlyTyrProGluGlyGlyIleAspAlaCysGlnGlyAspSer 352
QY 733 GGAGTTCACATCATGTGCGGGAATAAG-----AAAGGGCGCTGGACATCTGGCTGGT 783
DB 353 GlyGlyProPheValCysGluAspSerIleSerGlyThrSerArgTyrpArgLeuCysGly 372
QY 784 CTGACTTCCTGGGTTTGGGCTGTGGTCGAGGCTGGAGAACATGTGAGGAAAGTGCAT 843
DB 373 IleValSerTrpGlyThrGlyCysAlaLeu-----AlaArgLys----- 385
QY 844 CAAGGATCCCTGGGATCTTCACACACATATTAGTAAGTGCTTCTCTGGATCCACGAACAC 903
DB 386 -----ProGlyValTyrThrLysValThrAspPheArgGluTrpIlePheLysAla 402
QY 904 ATCCAAACT 912
DB 403 IleLysThr 405

RESULT 12
HEPS_HUMAN
ID HEPS_HUMAN STANDARD; PRT; 417 AA.
AC P05981;
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serine protease hepsin (EC 3.4.21.-) (transmembrane protease, serine
DE 1).
GN HPN OR TMPRSS1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=88209431; PubMed=2835076;
RA Leytus S.P., Loeb K.R., Hagen F.S., Kurachi K., Davie E.W.;
RT "A novel trypsin-like serine protease (hepsin) with a putative
RT transmembrane domain expressed by human liver and hepatoma cells.";
RL Biochemistry 27:1067-1074(1988).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=91358502; PubMed=1885621;
RA Tsuji A., Torres-Rosado A., Arai T., le Beau M.M., Lemons R.S.,
RA Chou S.H., Kurachi K.;
RT "Hepsin, a cell membrane-associated protease. Characterization,
RT tissue distribution, and gene localization.";
RL J. Biol. Chem. 266:16948-16953(1991).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=93348237; PubMed=8346233;
RA Torres-Rosado A., O'Shea K.S., Tsuji A., Chou S.H., Kurachi K.;
RT "Hepsin, a putative cell-surface serine protease, is required for
RT mammalian cell growth.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:7181-7187(1993).
CC -1- FUNCTION: PLAYS AN ESSENTIAL ROLE IN CELL GROWTH AND MAINTENANCE
CC OF CELL MORPHOLOGY.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
CC
```

```
CC -1- TISSUE SPECIFICITY: PRESENT IN MOST TISSUES, WITH THE HIGHEST
CC LEVEL IN LIVER.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use, by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M18930; AAA36013.1; -
CC EMBL; X07732; CAA30558.1; -
CC EMBL; X07002; CAA30058.1; -
CC PIR; S00845; S00845.
CC HSP; P00763; LDPO.
CC MEROPS; S01.224; -.
CC Genew; HGNC:5155; HPN.
CC MIM; 142440; -.
CC InterPro; IPR001314; Chymotrypsin.
CC InterPro; IPR001254; Ser_protease_Try.
CC InterPro; IPR001190; Srcr_receptor.
CC Pfam; PF00089; trypsin; 1.
CC PRINTS; PR00722; CHYMOTRYPSIN.
CC SMART; SM00202; SR; 1.
CC SMART; SM00202; Tryp_SPC; 1.
CC PROSITE; PS0240; TRYPSIN_DOM; 1.
CC PROSITE; PS00134; TRYPSIN_HIS; 1.
CC PROSITE; PS00135; TRYPSIN_SER; 1.
CC Hydrolase; Serine protease; Transmembrane; Signal-anchor.
CC CHAIN 1 162 NON-CATALYTIC CHAIN (POTENTIAL).
CC CHAIN 163 417 CATALYTIC CHAIN (POTENTIAL).
CC DOMAIN 1 17 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 18 44 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
CC TRANSMEM (POTENTIAL).
CC DOMAIN 45 417 EXTRACELLULAR (POTENTIAL).
CC DOMAIN 163 417 SERINE PROTEASE.
CC ACT_SITE 203 203 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT_SITE 257 257 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC ACT_SITE 353 353 CHARGE RELAY SYSTEM (BY SIMILARITY).
CC DISULFID 153 277 INTERCHAIN (BY SIMILARITY).
CC DISULFID 188 204 BY SIMILARITY.
CC DISULFID 322 338 BY SIMILARITY.
CC DISULFID 349 381 BY SIMILARITY.
CC CARBOHYD 112 112 N-LINKED (GLCNAC...) (POTENTIAL).
CC SEQUENCE 417 AA; 45011 MW; B2086FF6E1E551D7 CRC64;

Alignment Scores:
Pred. No.: 3.04e-33 Length: 417
Score: 440.00 Matches: 96
Percent Similarity: 49.82% Conservative: 42
Best Local Similarity: 34.66% Mismatches: 107
Query Watch: 26.38% Indels: 32
DB: 1 Gaps: 8

US-09-735-713a-1 (1-921) x HEPS_HUMAN (1-417)
QY 109 TGTGGCAGAGCTGGTTAAGGTACAGCCTTGGGAATTTTAAACATTTTCAGTCGCATT 168
DB 153 CysGlyArgArgLysLeuProVal-----AspArgile 163
QY 169 CTGTGAGGAAGCCCAAGTGGAGAAGGTTCTTCCTCCCTGGCAGGTATCTCTGAACAAAGG 228
DB 164 ValGlyGlyArgAspThrSerLeuGlyArgTyrpTrpTrpValSerLeuArgTyrAsp 183
QY 229 CAGAAGCATATTGTGGAGGAAGTCTCTCACCACACATGGTGGTATCAGCGCGCTCAC 288
DB 184 GlyAlaHisLeuCysGlyGlySerLeuLeuSerGlyAspTyrpValLeuThrAlaHis 203
QY 289 TGCATTGCAACAGAACAC---ATTGTGTCTACTTTGAATGTTACGTGGAGAGTATGAC 345
DB 204 CysPheProGluArgAsnArgValLeuSerArgTyrpArgValPheAlaGlyAla----- 221
```

QY 346 TTAAGCCAGACAGACCAGGAGAGCAAACTCTCACTATTGAACTGTCAATCATCATACATCCA 405  
Db 222 ValAlaGlnAlaSerProHisGlyLeuGlnLeuGlyValGlnAlaValAlaValTyrHisGly 241  
QY 406 CATTTC-----TCCACCAAGAACAACCAATGGACTATGATATTTGCCCTTTTG 450  
Db 242 GlyTyrLeuProPheArgAppProAsnSerGluGluAsnSerAsnAspIleAlaLeuVal 261  
QY 451 AAGATGGCTGGAGCCTCCAAATTTGGCCACTTTGTGGGCCCAATGATCTCTTCCAGAGCTG 510  
Db 262 HisLeuSerSerProLeuProLeuThrGluTyrIleGlnProValCysLeuProAlaAla 281  
QY 511 CGGAGCAATTTGAGGCTGTTTATTGTACAACTCCAGGCTGGGGCCGCTTAACATGAA 570  
Db 282 GlyGlnAlaLeuValaspGlyLysIleCysThrValThrGlyTrpGlyAsnThrGlnTyr 301  
QY 571 GGTGGCTCTCTCAAGTCTTTCAGGAAGTGAATCTGCCTATTGTTGACCTGGGAAGAG 630  
Db 302 TyrGlyGlnGlnAlaGlyValLeuGlnGluAlaArgValProIleIleSerAsnAspVal 321  
QY 631 TGTGTGGCAGCTCTGTTAACACTAAGAGGCCCATCATGCTGGGAAGAGCTTTCTTTGCACA 690  
Db 322 CysAsnGlyAlaAspPhe---TyrGlyAsnGlnIleLysProLysMetPhe---CysAla 339  
QY 691 GGTTCCTCTGATGAGGAGCAGCATGTCAGGAGATTCAGGAGTTCAGGAGTTCACATGTC 750  
Db 340 GlyTyrProGluGlyGlyIleAspAlaCysGlnGlyAspSerGlyGlyProPheValCys 359  
QY 751 CGG-----AATAAGAAGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 801  
Db 360 GluAspSerIleSerArgThrProArgTyrArgLeuCysGlyIleValSerTrpGlyThr 379  
QY 802 GGCTGTGCTGAGCTGGAGAAACAATGTGAGGAAAGTATCAAGATCCCTCGGAGTAC 861  
Db 380 GlyCysAlaLeuAlaGlnLys-----ProGlyVal 389  
QY 862 TTCACAGACATTAAGTAAGTCTTCTGATCCAGACACATCCAACT 912  
Db 390 TyrThrLysValSerAspPheArgGluTrpIlePheGlnAlaIleLysThr 406  
RESULT 13  
HEPS\_RAT  
ID HEPS\_RAT STANDARD; PRT; 416 AA.  
AC Q05511;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Serine protease hepsin (EC 3.4.21.-).  
GN HPN.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=93305733; PubMed=8318546;  
RA Farley D., Reymond F., Nick H.;  
RT "Cloning and sequence analysis of rat hepsin, a cell surface serine  
RT proteinase";  
RL Biochim. Biophys. Acta 1173:350-352(1993).  
CC -!- FUNCTION: PLAYS AN ESSENTIAL ROLE IN CELL GROWTH AND MAINTENANCE  
CC OF CELL MORPHOLOGY.  
CC -!- SUBCELLULAR LOCATION: Type II membrane protein.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC -----  
CC EMBL; X70900; CAA50256.1; -.  
DR PIR; S32013; S32013.  
DR PIR; S33777; S33777.  
DR HSP; P00763; IDPO.  
DR MEROPS; S01.224; -.  
DR InterPro; IPR001314; Chymotrypsin.  
DR InterPro; IPR01254; Ser\_protease\_Try.  
DR InterPro; IPR001190; Srcr\_receptor.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYMOTRYPsin.  
DR SMART; SM00202; SR; 1.  
DR SMART; SM00020; Tryp\_SPC; 1.  
DR PROSITE; PS0240; TRYPsin\_DOM; 1.  
DR PROSITE; PS00134; TRYPsin\_HIS; 1.  
DR PROSITE; PS00135; TRYPsin\_SER; 1.  
KW Hydrolase; Serine protease; Transmembrane; Signal-anchor.  
FT CHAIN 1 161 NON-CATALYTIC CHAIN (POTENTIAL).  
FT CHAIN 162 416 CATALYTIC CHAIN (POTENTIAL).  
FT DOMAIN 1 16 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 17 43 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
(POTENTIAL).  
FT DOMAIN 44 416 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 162 416 SERINE PROTEASE.  
FT ACT\_SITE 202 202 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 256 256 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT ACT\_SITE 352 352 CHARGE RELAY SYSTEM (BY SIMILARITY).  
FT DISULFID 152 276 INTERCHAIN (BY SIMILARITY).  
FT DISULFID 187 203 BY SIMILARITY.  
FT DISULFID 321 337 BY SIMILARITY.  
FT DISULFID 348 380 BY SIMILARITY.  
FT CARBOHYD 111 111 N-LINKED (GLCNAC... (POTENTIAL).  
SQ SEQUENCE 416 AA; 44926 MW; E5A9F8FA9550E180 CRC64;  
Alignment Scores:  
Pred. No.: 4.68e-33 Length: 416  
Score: 438.00 Matches: 99  
Percent Similarity: 49.47% Conservative: 41  
Best Local Similarity: 34.98% Mismatches: 99  
Query Match: 26.26% Indels: 44  
DB: 1 Gaps: 10  
US-09-735-713a-1 (1-921) x HEPS\_RAT (1-416)  
QY 109 TGTGGCAGAGTCTGTTAAGGTACAGCCTTGGAATTATTTTAAACATTTTCAGTCGCATT 168  
Db 152 CysGlyArgArgLysLeuProVal-----AspArgile 162  
QY 169 CTGGAGGAGCAACTGGAGAGGTTCTCTATCCCTGGCAGGTATCTCTGAAACAAAGG 228  
Db 163 ValGlyGlyGlnAspSerSerLeuGlyArgTyrProTyrGlnValSerLeuArgTyrAsp 182  
QY 229 CAGAAGCATATTTGTGGAGAGCATCTCTCACCACAGTGGGTGATACACGGGGCTCAC 288  
Db 183 GlyThrHisLeuCysGlyGlySerLeuLeuSerGlyAspTrpValLeuThrAlaAlaHis 202  
QY 289 TCATTGCAACACAGAAC---ATTGTTCTACTTTCAATGTTACTCTCTGGAGAGTATGAC 345  
Db 203 CysPheProGluArgAsnArgValLeuSerArgTyrArgValPheAlaGlyAla----- 220  
QY 346 TTAAGCCAGACAGACCCAGGAGAGCAAACTCTCACTATTGAACTGTCAATCATCATACATCCA 405  
Db 221 ValAlaArgThrSerProHisAlaValGlnLeuGlyValGlnAlaValIleTyrHisGly 240  
QY 406 CATTTC-----TCCACCAAGAACAACCAATGGACTATGATATTTGCCCTTTTG 450  
Db 241 GlyTyrLeuProPheArgAppProThrIleAspGluAsnSerAsnAspIleAlaLeuVal 260  
QY 451 AAGATGGCTGGAGCCTCCAAATTTGGCCACTTTGTGGGCCCAATGATCTCTTCCAGAGCTG 510  
Db 261 HisLeuSerSerSerLeuProLeuThrGluTyrIleGlnProValCysLeuProAlaAla 280

```
Qy 511 CGGAGCAATTTAGCGCTGTTTATTGTTACAACTGACGAGCTGGGCGCTTAACCTGAA 570
Db 281 GlyGlnAlaLeuValAspGlyValCysThrValThrGlyTrpGlyAsnThrClnPhe 300
Qy 571 GGTGGCTGCTCTCACAAGCTTTCAGAGAAAGTAATCTGCTATTTTACCTGAGGAGAG 630
Db 301 TyrGlyGlnGlnAlaValLeuGlnGlnAlaArgValProIleSerAsnGluVal 320
Qy 631 TGTGTGGCAGCTCTCTTAACACTAAGAGGCC-----ATCAGTGGG 672
Db 321 Cys-----AsnSerProAspPheTyrGlyAsnGlnIleLysPro 333
Qy 673 AAGACCTTTCTTTCACAGGTTTCTGATGAGGAGAGACGCGATGTCAGGAGATTC 732
Db 334 LysMetPhe---CysAlaGlyTyrProGluGlyIleAspAlaCysGlnGlyAspSer 352
Qy 733 GGAGTTCACATCNGTGGCGGAATAAG-----AAAGGGCGCTGACTCTGGCTGGT 783
Db 353 GlyGlyHisPheValCysGluAspArgIleSerGlyThrSerArgTrpArgLeuCysGly 372
Qy 784 GTGACTTCTGCTGGGTTTGGCTGTGCTGAGGCTGAGAGAAACAATGTGAGGAAAGTGT 843
Db 373 IleValSerTrpGlyThrGlyCysAlaLeu-----AlaArgLys----- 385
Qy 844 CAAGGATCCCTGGGATCTCACAGACATTAAGTAAAGTCTTCTGATGATCCAGCAAC 903
Db 386 -----ProGlyValTyrThrLysValIleAspPheArgGluTrpIlePheGlnAla 402
Qy 904 ATCCAACT 912
Db 403 IleLysThr 405

RESULT 14
TMS5_HUMAN STANDARD; PRT; 457 AA.
ID TMS5_HUMAN Q9H3S3;
AC 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Transmembrane protease, serine 5 (EC 3.4.21.-) (Spinesin).
GN TMRSS5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX PubMed-11741986;
RA Yamaguchi N., Okui A., Yamada T., Nakazato H., Mitsui S.;
RT "Spinesin/TMRSS5, a novel transmembrane serine protease, cloned from
RT human spinal cord."
RL J. Biol. Chem. 277:6806-6812(2002).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: Brain-specific. Predominantly expressed in
CC neurons, in their axons, and at the synapses of motoneurons in the
CC spinal cord.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -1- SIMILARITY: CONTAINS 1 SRCR DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB028140; BAB20375.1; -
DR HSSP; P00763; 1DPO.
DR Genew; HGNC:14908; TMRSS5.
DR MIM; 606751; -
DR MEROPS; S01.313; -
```

```
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR001190; Srcr_receptor.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00202; SR; 1.
DR SMART; SM00202; Tryp_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; 1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
DR PROSITE; PS00420; SRCR_1; FALSE_NEG.
DR PROSITE; PS0287; SRCR_2; FALSE_NEG.
DR Hydrolase; Serine protease; Transmembrane; Signal-anchor;
KW Glycoprotein.
FT DOMAIN 1 49 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 50 70 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT DOMAIN 71 457 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 112 207 SRCR.
FT DOMAIN 218 457 SERINE PROTEASE.
FT ACT_SITE 258 258 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 308 308 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 405 405 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT SITE 217 218 CLEAVAGE (POTENTIAL).
FT DISULFID 135 196 BY SIMILARITY.
FT DISULFID 148 206 BY SIMILARITY.
FT DISULFID 209 328 BY SIMILARITY.
FT DISULFID 243 259 BY SIMILARITY.
FT DISULFID 374 390 BY SIMILARITY.
FT DISULFID 401 429 BY SIMILARITY.
FT CARBOHYD 163 163 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 170 170 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 195 195 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 319 319 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 375 375 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 457 AA; 49574 MW; 64406AB4985A2651 CRC64;

Alignment Scores:
Pred. No.: 4.78e-33 Length: 457
Score: 438.00 Matches: 93
Percent Similarity: 53.91% Conservative: 45
Best Local Similarity: 36.33% Mismatches: 94
Query Match: 26.26% Indels: 24
DB: 1 Gaps: 8

US-09-735-713A-1 (1-921) x TMS5_HUMAN (1-457)
Qy 160 AGTCGCAATCTTGGAGAGCAAGTGGAGAGGGTCTCTATCCCTGCGAGGATCTCTG 219
Db 216 SerArgIleValGlyGlyGlnSerValAlaProGlyArgTrpProTrpGlnAlaSerVal 235
Qy 220 AAACAAAGGCAGACATATTTGTGGAGGAGCATCTCTCACCACAGCTGGTGATCAGG 279
Db 236 AlaLeuGlyPheArgHisThrCysGlyGlySerValLeuAlaProArgTrpValValThr 255
Qy 280 CGCGCTACTGCAATGCAACAGAACATTT-----GTGCTCTACTTTGAATGTTACTGCT 333
Db 256 AlaAlaHisCysMetHisSerPheArgLeuAlaArgLeuSerSerTrpArgValHisAla 275
Qy 334 GGAGAGTATGACTTAAGCCAGACACACCCAGGAGAGCAAACTCTCATTGGAACACTGTC 393
Db 276 GlyLeuValSerHisSerAlaValArgProHisGlnGlyAlaLeu--ValGluArgIle 294
Qy 394 ATCATATCATCCATTTCTCCACCAAGAACCAATGGACTATGATATGCGCTTTTGAAG 453
Db 295 IleProHisProLeuTyrSerAlaGlnAsn---HisAspTyrAspValAlaLeuLeuArg 313
Qy 454 ATGGCTGAGGCTTCCAATTTGGCCACTTTGTGGGCCCATATGTTCTCCAGAGCTGGCG 513
Db 314 LeuGlnThrAlaLeuAsnPheSerAspThrValGlyAlaValCysLeuProAlaLysGlu 333
Qy 514 GAGCAATTTGAGCGTGGTATTTATTTGTTACAACTGCAGGCTGGGCGCGCTTAACCTGAGGT 573
```

Db 334 GlnHisPheProLysGlySerArgCysTrpValSerGlyTrpGlyHisThrHisProSer 353

QY 574 GGCGTCTCTC---TCACAGCTTTCAGGAGTGAATCTGCTATTTGACCTGGGAAGAG 630

Db 354 HisThrTyrSerSerAspMetLeuGlnAspThrValValProLeuPheSerThrGlnLeu 373

QY 631 TGTGTGGCAGCTCTGTAAACATAAGAGAGCCCATCAGTGGG-----AAGACCTTT 681

Db 374 CysAsnSerSerCysVal-----TyrSerGlyAlaLeuThrProArgMet 388

QY 682 CTTTGGACAGCTTTCCTGATGAGGAGAGACGATGTCAGGAGATTCAGGAGCTTCA 741

Db 389 LeuGlyAlaGlyTyrLeuAspGlyArgAlaAspAlaCysGlnGlyAspSerGlyGlyPro 408

QY 742 CTCATGTGCGGGAATAAGAAAGGGCGCTGACTCTGCTGGTGTGACTCTCCTGGGCTTG 801

Db 409 LeuValCysProAsp---GlyAspThrTrpArgLeuValGlyValSerTrpGlyArg 427

QY 802 GCGTGTGTCGAGCTGGAGGAAACAATGTGAGGAAAGTATCAAGGATCCCTGGGATC 861

Db 428 AlaCys-----AlaGluProAsnHisProGlyVal 437

QY 862 TTCACACACATTAGTAAGTGCCTTCTCGATCCAGCAACACATCCAA 909

Db 438 TyrAlaLysValAlaGluPheLeuAspTrpPheHisAspThrAlaGln 453

RESULT 15

FA12\_BOVIN STANDARD; PRT; 593 AA.

AC P98140;

DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DE 15-JUN-2002 (Rel. 41, Last annotation update)

DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)

DE (HAF) (Fragment).

GN Fl2.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;

OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RX MEDLINE=94242782; PubMed=8186251;

RA Shibuya Y., Samba U., Okabe H., Kambara T., Yamamoto T.;

RT "Primary structure of bovine Hageman factor (blood coagulation factor XII); comparison with human and guinea pig molecules.";

RL Biochim. Biophys. Acta 1206:63-70(1994).

RN [2]

RP SEQUENCE OF 10-21; 350-364 AND 525-550.

RX MEDLINE=77182112; PubMed=861210;

RA Fujikawa K., Walsh A.K., Davie W.E.;

RT "Isolation and characterization of bovine factor XII (Hageman factor).";

RL Biochemistry 16:2270-2278(1977).

CC -1- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE GENERATION OF BRADYKININ AND ANGIOTENSIN.

CC -1- CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor VII to form factor VIIa and factor XI to form factor Xla.

CC -1- PTM: O- AND N-GLYCOSYLATED (BY SIMILARITY).

CC -1- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XI FIRST TO ALPHA-FACTOR Xla AND THEN TO BETA-FACTOR Xlaa. ALPHA-FACTOR Xlaa ACTIVATES FACTOR XI TO FACTOR Xla. BOVINE FACTOR XII IS CLEAVED ONLY TO ALPHA-FACTOR Xlaa AS IT LACKS THE TRYPSIN/KALLIKREIN CLEAVAGE SITE.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.

CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; S70164; AAB30804.2; --

DR HSP; P00763; IDPO.

DR MEROPS; S01.211; --

DR InterPro; IPR000561; EGF-like.

DR InterPro; IPR000562; FN\_Type\_II.

DR InterPro; IPR000083; Fibrinctn1.

DR InterPro; IPR000001; Kringle.

DR InterPro; IPR001254; Ser\_protease\_Try.

DR Pfam; PF00008; EGF; 2.

DR Pfam; PF00039; fn1; 1.

DR Pfam; PF00040; fn2; 1.

DR Pfam; PF00051; kringle; 1.

DR Pfam; PF00089; trypsin; 1.

DR ProDom; PD000395; Kringle; 1.

DR ProDom; PD000995; FN\_Type\_II; 1.

DR SMART; SM00181; EGF; 2.

DR SMART; SM00058; FN1; 1.

DR SMART; SM00059; FN2; 1.

DR SMART; SM00130; KR; 1.

DR SMART; SM00020; Tryp\_Spc; 1.

DR PROSITE; PS00022; EGF\_1; 2.

DR PROSITE; PS01186; EGF\_2; FALSE\_NEG.

DR PROSITE; PS01253; FIBRONECTIN\_1; 1.

DR PROSITE; PS00023; FIBRONECTIN\_2; 1.

DR PROSITE; PS00021; KRINGLE\_1; 1.

DR PROSITE; PS50070; KRINGLE\_2; 1.

DR PROSITE; PS50240; TRYPSIN\_DOM; 1.

DR PROSITE; PS00134; TRYPSIN\_HIS; 1.

DR PROSITE; PS00135; TRYPSIN\_SER; 1.

KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;

KW Hydrolase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.

FT NON\_TER 1

FT SIGNAL <1 9

FT CHAIN 10 349

FT CHAIN 350 593

FT DOMAIN 37 78

FT DOMAIN 84 121

FT DOMAIN 123 163

FT DOMAIN 164 200

FT DOMAIN 207 287

FT DOMAIN 297 333

FT DOMAIN 350 593

FT ACT\_SITE 389

FT ACT\_SITE 438

FT ACT\_SITE 541

FT DISULFID 88 100

FT DISULFID 94 109

FT DISULFID 111 120

FT DISULFID 125 153

FT DISULFID 151 160

FT DISULFID 168 179

FT DISULFID 173 188

FT DISULFID 190 199

FT DISULFID 207 287

FT DISULFID 230 269

FT DISULFID 258 282

FT DISULFID 336 463

FT DISULFID 374 390

FT DISULFID 382 452

FT DISULFID 413 416

FT DISULFID 479 547

FT DISULFID 510 526

FT DISULFID 537 568

FT DISULFID 537 568

Search completed: March 25, 2003, 06:58:49  
Job time : 25.5 secs

```
FT CARBOHYD 99 99 O-LINKED (FUC) (BY SIMILARITY).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 410 410 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 593 AA; 65148 MW; 721592BA792BD61F CRC64;

Alignment Scores:
Pred. NO.: 1.48e-32 Length: 593
Score: 433.00 Matches: 101
Percent Similarity: 52.14% Conservative: 45
Best Local Similarity: 36.07% Mismatches: 94
Query Match: 25.96% Indels: 40
DB: 1 Gaps: 12

US-09-735-713A-1 (1-921) x FA12_BOVIN (1-593)

Qy 109 TGTGGGACAGCTCGTTAAGGTACAGCCTTGGAAATATTTTAAACATTTTCAGTCGCATT 168
Db 336 CysGlyGlnArgLeuArgLys-----TrpLeuSerSerLeuAsnArgVal 350
Qy 169 CTGGAGGAGCCAAAGTGGAGAGGTTCTCTATCCCTCGCAGGTTATCTCTGAACAAAGG 228
Db 351 ValGlyGlyLeuValAlaLeuProGlyAlaHisProIleAlaAlaLeuTrpAsp 370
Qy 229 CAGAAGCATATTTGTGGAGGAGCATCGTCTCACCACATGGGTGATCACGGCGCTCAC 288
Db 371 Gln---HisPheCysAlaGlySerLeuIleAlaProCysTrpValLeuThrAlaAlaHis 389
Qy 289 TGCATTGCAACACAGAACATTTGTCTACTTTTGAATGTACTGTCTGGAGAGATGACTTA 348
Db 390 CysLeuGlnAsnArgProAlaProIysGluLeuThrValValLeuGlyGlnAspArgHis 409
Qy 349 AGCCAGACAGACCAGGAGAGCAACTCTCCTACTATTGAACTGTCATCATCATCCACAT 408
Db 410 AsnGlnSerCysGluGlnCysGlnThrLeuAlaValArgAspTyrArgLeuHisGluAla 429
Qy 409 TTCTCCACCAAGAACCAATGGACTAT-----GATATTGCCCTTTTGAAGATGGCTGGA 462
Db 430 PheSer-----ProIleThrTyrGlnHisAspLeuAlaLeuValArgLeuGlnGlu 446
Qy 463 GCCTTCCAA-----TTTGGCCAC-----TTTGGGGCCCATATGCTTCCAGAG 507
Db 447 SerAlaAspGlyCysCysAlaHisProSerProPheValGlnProValCysLeuProSer 466
Qy 508 -----CTGCGGAGCAATTTGAGCGTGTATTGTTTGTACAACCTGCAGGCTGGGGCGC 561
Db 467 ThrAlaAlaArgProAlaGluSerGluAlaAlaValCysGluValAlaGlyTrpGlyHis 486
Qy 562 TTAAGTGAAGTGGCGTCTCTCACAAGTCTTGCAAGGAAGTGAATCTGCCTATTTTGACC 621
Db 487 GlnPheGluGlyGlyGlyTyrSerSerPheLeuGlnGluAlaGlnValProLeuIleAsp 506
Qy 622 TGGNAGAGTGTGTGGCAGCTCTGTTACACTAAAGAGGCCCATCAGTGGGAAGACCTTT 681
Db 507 ProGlnArgCysSerAlaPro-----AspValHisGlyAlaAlaPhe 520
Qy 682 -----CTTTCACAGGTTTCTCATGAGGAGAGAGAGCGCATGTCAGGAGAT 729
Db 521 ThrGlnGlyMetLeuGlyAlaGlyPheLeuGlyGlyThrAspAlaCysGlnGlyAsp 540
Qy 730 TCAGAGGTTTCATCATGTCGCGGAATAAG-----AAAGGGCGCTGGACTCTGGCTGCT 783
Db 541 SerGlyGlyProLeuValCysGluAspGluThrProGluArgGlnLeuIleLeuArgGly 560
Qy 784 GTGACTTCTGGGTTGGCTGTGGTCGAGGCTGGAGAAACAATGTGAGGAAAGTGAT 843
Db 561 IleValSerTrpGlySerGlyCysGly-----AsnArgLeuLys----- 573
Qy 844 CAAGGATCCCTGGGATCTTCACACACATAGTAAGTGTCTTCTGGATCCACGAACAC 903
Db 574 -----ProGlyValTyrThrAspValAlaAsnTyrLeuAlaTrpIleArgGluHis 590
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: March 25, 2003, 06:51:58 ; Search time 90.5 seconds  
(without alignments)  
4193.799 Million cell updates/sec

Title: US-09-735-713A-1

Perfect score: 1668

Sequence: 1 atgagtctcaaatgcttat.....acatccaaactgtaactaa 921

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame-n2p.model -DEV=xlh  
-Q/cgn2\_1/USPTO.spool/US09735713/runat\_18032003\_124354\_23129/app\_query.fasta\_1.1095  
-DB=SPTRMBL\_21 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptio -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09735713\_CGN\_1\_1\_57\_@runat\_18032003\_124354\_23129 -NCPU=6 -ICPU=3  
-NO\_XLPAY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	690.5	41.4	1004	13	P79953	P79953 xenopus lae

	2	657	39.4	974	13	Q90WD8
	3	426	37.5	1524	13	Q91674
	4	496	29.7	277	5	Q96899
	5	486	29.1	787	5	Q9VEY6
	6	480.5	28.8	799	11	Q9DBI0
	7	475.5	28.5	335	11	Q8VIF2
	8	473.5	28.4	624	11	Q9DAT3
	9	467.5	28.0	624	11	Q91V47
	10	462.5	27.7	467	5	Q967X8
	11	462	27.7	453	11	Q8VDE0
	12	458	27.5	638	11	Q8R0P5
	13	450	27.0	339	11	Q99L44
	14	446	26.7	643	6	Q97506
	15	441.5	26.5	433	13	Q90VK1
	16	436.5	26.2	317	13	Q9DGR3
	17	430	25.8	767	13	Q9DGR2
	18	429.5	25.7	310	11	Q91XC4
	19	428.5	25.7	581	5	Q96015
	20	428.5	25.7	1047	5	Q9VZH2
	21	427.5	25.6	387	5	Q9XY57
	22	426	25.5	624	6	Q95ME7
	23	424.5	25.4	310	11	Q9QYZ9
	24	423.5	25.4	616	6	Q97507
	25	422.5	25.3	267	5	Q9BK47
	26	422.5	25.3	456	6	Q9TTR0
	27	422	25.3	314	5	Q9VR15
	28	421	25.2	273	6	Q9XSM1
	29	421	25.2	855	11	Q9JJI7
	30	420.5	25.2	1047	5	Q24019
	31	420	25.2	321	4	Q96RZ8
	32	419	25.1	256	5	Q9XYV1
	33	418.5	25.1	325	5	Q15944
	34	418.5	25.1	329	6	Q9GL10
	35	417	25.0	264	11	Q9D7P8
	36	417	25.0	264	11	Q9ER05
	37	414.5	24.9	300	4	Q96EF3
	38	414.5	24.9	573	5	Q9V516
	39	414	24.8	254	5	Q8T637
	40	414	24.8	264	11	Q9EQZ8
	41	413.5	24.8	268	5	Q46151
	42	413	24.8	263	11	Q9DC86
	43	413	24.8	405	4	Q96E86
	44	412	24.7	263	11	Q9CR35
	45	409.5	24.6	468	5	Q9U0G3

#### ALIGNMENTS

RESULT 1

P79953  
ID P79953 PRELIMINARY; PRT: 1004 AA.  
AC P79953;  
DT 01-MAY-1997 (TREMBLrel. 03, Created)  
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Oviductin.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;  
OC Xenopodinae; Xenopus.  
NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE:99184825; PubMed:10084976;  
RA Lindsay L.L., Wieduwilt M.J., Hedrick J.L.;  
RT "Oviductin, the Xenopus laevis oviductal protease that processes egg  
RT envelope glycoprotein gp43, increases sperm binding to envelopes, and  
RT is translated as part of an unusual mosaic protein composed of two  
RT protease and several CUB domains.";  
RL Biol. Reprod. 60:989-995(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Yang J.C., Lindsay L.L., Hedrick J.L.;

```
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 3 CUB DOMAINS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; U81291; AAB53972.1; -.
DR HSP; P00763; IDPO.
DR MEROPS; S01.240; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR000859; CUB_domain.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00431; CUB; 2.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00042; CUB; 3.
DR SMART; SM00020; Tryp_SPC; 2.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS0240; TRYPSIN_DOM; 2.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 1004 AA; 110612 MW; 5ECACE265E6433CE CRC64;

Alignment Scores:
Pred. No.: 1-22e-62 Length: 1004
Score: 690.50 Matches: 132
Percent Similarity: 60.33% Conservative: 49
Best Local Similarity: 44.00% Mismatches: 110
Query Match: 41.40% Indels: 9
DB: 13 Gaps: 4

US-09-735-713a-1 (1-921) x P79953 (1-1004)
QY 22 ACAGGAACAGCTGATTTTACTAGTAGTCTTTTGAACRAGGTAAATCTGCA 81
DB 3 ThrArgAsn-----LeuLeuLeuGlySerIleLeuLeuSerLeuAlaValLysGly 19
QY 82 RCTCTTTCGTCCTCCCAAGCTCCACGTCTGTGGCAGAGT---CTGTTAAGGTACAGCT 138
DB 20 AspproGlyProHisArgGlyAlaArgCysGlyValSerProLeuGlySerAlaThrGlu 39
QY 139 TGAATATTATTAACATTTTCAGTCGCATCTTGTGGAGGAGCAAGTGGAGAGGGTTCC 198
DB 40 LeuAsnTyr-----LeuSerArgIleValGlyGlyArgGluSerLysLysGlyGln 56
QY 199 TATCCCTGCGAGTATCTCTGAAACAAAGCAGACAGCATATTGTGGAGAGACATCGTC 258
DB 57 HisProTrpThrValSerLeuLysArgAsnGlyLysHisPheCysGlyThrLeuVal 76
QY 259 TCACCACAGTGGGTGATCATCGCGGCTCACTGCATTTGCAACAGAACATTTGCTCTACT 318
DB 77 SerHisCysHisValLeuThrAlaAlaHisCysLeuLeuAspArgAsnValLysLeuTyr 96
QY 319 TTGAATGTTTACTGCTGGAGAGATGACTTTAAGCCAGACAGCCAGAGCAAACTCTC 378
DB 97 MetArgValTyrIleGlyGluTyrAspGlnIleLeuLysGluGluThrGluGlnMetPhe 116
QY 379 ACTATTGAACCTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 438
DB 117 ArgValIleGluIlePheLysHisProAsnPheAsnGlnSerGlnProMetAsnTyrAsp 136
QY 439 ATTGCCCTTTTGAAGATGCTGAGGCTTCCAAATTTGGCCACTTTTGTGGGGCCCATATGT 498
DB 137 ValAlaValLeuLeuLeuAspGlySerValThrPheAspGluAsnIleGlnProAlaCys 156
QY 499 CTTCCAGAGCTGCGGAGCAATTTGAGGCTGTTTATTTTGTACAACTGCAGCTGGGCG 558
DB 157 LeuProAsnProAspValPheGluProGlyAspLeuCysValThrLeuGlyTrpGly 176
QY 559 CGCTTAAGTGAAGTGGCTCTCTCATCAAGTCTTGCAGGAAGTGAATCTGCTATTTTG 618
DB 177 HisLeuThrGluAsnGlyIleLeuProValValLeuGlnGluValTyrLeuProIleVal 196
QY 619 ACCTGGGAAGAGTGTGTGCACCTCTGTAACTAAAGAGGCCCATCAGTGGGAGACC 678
```

```
Db 197 AspleuSerSerCysLeuHisValMetSerAlaLeuLysGlyThrValValSerSerTyr 216
QY 679 TTTCTTTGCACAGTTTCTCTGATGAGGAGAGAGCATGTCTCAGGAGATTCAGGAGGT 738
Db 217 IleValCysAlaGlyPheProGluGlyLysAspAlaCysGlnGlyAspSerGlyGly 236
QY 739 TCACATCATGTGCGGAATAAGAAAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 798
Db 237 ProLeuLeuCysGlnArgHisGlySerTrpValLeuHisGlyLeuThrSerTrpGly 256
QY 799 TTGGGCTGTGCTGAGGCTGGAGAAACAATGTG-----AGGAAAAAGTGCATCAAGATCC 852
Db 257 MetGlyCysGlyArgSerTrpLysAsnValPheLeuProHisAsnArgLysGlySer 276
QY 853 CTGGGATCTTCACACACATTAAGTGCCTTCTGATCCACGACATCCAAACT 912
Db 277 ProGlyIlePheThrAspIleGlnLysLeuLeuGlyTrpValSerSerGlnLeuAsnThr 296
RESULT 2
Q90WD8
ID Q90WD8 PRELIMINARY; PRT; 974 AA.
AC Q90WD8:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Oviductin.
OS Bufo japonicus (Japanese toad).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Bufo.
ON NCBI_TaxID=8387;
RX SEQUENCE FROM N.A.
RC TISSUE=OVIDUCTAL PARS RECTA;
RA Hiyoishi M., Takamune K., Mita K., Kubo H., Sugimoto Y., Katagiri C.;
RT "Oviductin, the oviductal protease that mediates gamete interaction by
RT affecting the vitelline envelope in Bufo japonicus: its molecular
RT cloning and analyses of expression and post-translational
RT activation."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB070367; BAB63372.1; -.
DR MEROPS; S01.240; -.
DR InterPro; IPR000859; CUB_domain.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00431; CUB; 2.
DR Pfam; PF00089; trypsin; 2.
DR PROSITE; PS01180; CUB; 2.
DR PROSITE; PS0240; TRYPSIN_DOM; 2.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 974 AA; 107647 MW; F19705A470465553 CRC64;
```

```
Alignment Scores:
Pred. No.: 3.65e-59 Length: 974
Score: 657.00 Matches: 116
Percent Similarity: 62.12% Conservative: 66
Best Local Similarity: 39.59% Mismatches: 103
Query Match: 39.39% Indels: 8
DB: 13 Gaps: 3

US-09-735-713A-1 (1-921) x Q90WD8 (1-974)
```

```
QY 34 CTGATTTTACTAGTAAGTCTTTTGAACRAGGTAAATCTGCARCTCTTTCGCTC 93
Db 12 LeuThrValMetIleGlyVal-----GlyArgGlyValThrAspSerPro 26
QY 94 CCAAAGCTCCAGTTGTGGCAGAGTCTGTTAAGGTACAGCCCTTGGAAATATTATTAAC 153
Db 27 GlyArgValSerArgCysGlyCiluArgProAlaAlaAsnThrSerValSerTyr---Gly 45
QY 154 ATTTTCAGTCGATCTTCTGGAGAGCCAAAGTGGAGAGGGTTCTATCCCTGGCAGGTA 213
Db 154 ATTTTCAGTCGATCTTCTGGAGAGCCAAAGTGGAGAGGGTTCTATCCCTGGCAGGTA 213
```



```

QY 604 AATCTGCCTATTTTACCTGGGAGAGTGTGTGGCAGCTCTGTAACTAAAGAGGCC 663
    |||||:||||:
Db 205 LysLeuProValValAspAsnGlyThrCysHisAlaValLeuGluProIleGlyHisPro 224
    |||||:||||:
QY 664 ATCAGTGGGAGACCTTTCTTCCACAGGTTTCTGTGATGGAGGAGAGCGCATGTCAG 723
    :||: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 225 ValLeuAspSerThrMetLeuCysAlaGlyPheProGluGlyGlyMetAspAlaCysGln 244
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 724 GGAGATTACAGAGGTTCACTCATGTGCCGAATAAGAAAGGGCCCTGGCTGGT 763
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 245 GlyAspSerGlyGlyPropheValCysArgArgSerGlyValIlePheLeuAlaGly 264
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 784 GTGACTTCTGGGTTTGGCTGTGTGCGAGCTGG---AGAAACAATGTGAGAAAGT 840
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 265 CysValSerTrpGlyLeuGlyCysGlyArgSerTrpGlyAlaLysGlnIleIleArgSer 284
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 841 GATCAAGATCCCTGGGCTTTCACAGACATTAAGTAAAGTCTTCTTGGATCCAGAA 900
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 285 GlnSerGlySerProAlaIlePheSerArgValSerValLeuAspPheLeuArgPro 304
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 901 CACATCCAAACTGGT 915
    |||||
Db 305 ProLysLeuThrGly 309
    |||||

RESULT 4
O96899 ID O96899 PRELIMINARY; PRT; 277 AA.
AC O96899;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Plasminogen activator SPA.
OS Scolopendra subspiniipes.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Myriapoda; Chilopoda;
OC Pleurostigmophora; Scolopendromorpha; Scolopendridae; Scolopendra.
OX NCBI_TaxID=55038;
RN [1]
RP SEQUENCE FROM N.A.
RA Sohn Y.-D., You W.K., Kim K.Y., Chung K.H., Park D.-H.;
RT "Cloning of a cDNA for a novel plasminogen activator from a Korean
RT centipede Scolopendra."
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
DR EMBL; U79521; AAD00320.1; -.
DR HSSP; P00763; 1DPO.
DR MEROPS; S01.122; -.
DR InterPro; IPR001314; Chymotrypsin.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; TRYPSIN_SPC; 1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 277 AA; 30172 MW; EC62F80C8CB246C CRC64;

Alignment Scores:
Pred. No.: 1.37e-42 Length: 277
Score: 496.00 Matches: 112
Percent Similarity: 53.74% Conservative: 46
Best Local Similarity: 38.10% Mismatches: 98
Query Match: 29.74% Indels: 38
DB: 5 Gaps: 12

US-09-735-713A-1 (1-921) x O96899 (1-277)
QY 28 AACAACTGATTTTACTACTAGGAATGCTTTTGAACRAGGTAAATCTGCARCTCTT 87
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 2 AsnSerPheThrIleLeuIle---ValThrTyrPhe----- 12

```

```

QY 88 TCGTCCCAAAAGCTCCCACTGGTGTGGCAGAGTCTGTGTTAAGGTACAGCTTGAATAT 147
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 13 SerLeuAlaPheGlySerArgCysGly-----IleLysAsnGlyPro-----Met 27
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 148 TTTAACTATTTTCAGTCGCATTTCTTGGAGGAGCCAAAGTGGAGAGGTTCTTATCCCTGG 207
    :||: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 28 LeuAspGluPheAsnArgIleValGlyGlyGluAlaAlaGluProGlyGluPheProTrp 47
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 208 CAGGTATCTCTGAAA-----CAAAGGCAGAACATATTTGTGGAGGAACATC 255
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 48 GlnIleSerLeuGlnValValSerTrpTyrGlySerTyrHisTyrCysGlyGlySerIle 67
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 256 GTCTACACAGCTGGGTGATCAGCGCGCTCACTGCATTGCAACACAGAACTTGTGTCT 315
    :||: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 68 LeuAspGluSerTrpValThrAlaAlaHisCysValGluGlyMetAsn---ProSer 86
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 316 ACTTTCAATCTTACTCTGAGAGTATGACTTAAGCCACACAGACACAGAGAGAACT 375
    |||: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 87 AspLeuArgIleLeuAlaGlyGluHisAsnPheLysLysGluAspGlyThrGluGlnTrp 106
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 376 CTCACATATTGAAACTCTCATCATACATTCACATTTCTCCACCAAGAAACCAATGGACTAT 435
    :||: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 107 GlnAspValIleAspIleIleMetHisLysAspTyr---ValTyrSerThrLeuGluAsn 125
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 436 GATATTGCCCTTTTGAAGTGGCTGGAGCTTCCAAATTT---GGCCACTTTGTGGGGCCC 492
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 126 AspIleAlaLeuLeuLysLeuAlaGluProLeuAspLeuThrProThrAlaValGlySer 145
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 493 ATATGCTTCCAGAGCTGCGGGAGCAATTTGAGGCTGGTTTATTTGTACAACTGCAGGC 552
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 146 IleCysLeuProSerGlnAsnGlnGluPheSerGlyHis---CysIleValThrGly 164
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 553 TGGGGCGCTTAACCTGAAGTGGCTCTCACAACTCTTGCGAGGAAGTGAATCTGCCT 612
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 165 TrpGlySerValArgGluGlyGlyAsnSerProAsnIleLeuGlnLysValSerValPro 184
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 613 ATTTTGACCTGGGAGAGTGTGTGGCAGCTCTGTTAACACTAAAGAGGCCCATCAGTGGG 672
    :||: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 185 LeuMetThrAspGluGluCysSerGluTyrTyrAsnIleVal----- 198
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 673 AGACCTTTCTTTCACAGCTTTTCTGATGGAGGAGAGACGCATCTCAGGAGATTCATCA 732
    |||: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 199 AspThrMetLeuCysAlaGlyTyrAlaGluGlyGlyLysAspAlaCysGlnGlyAspSer 218
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 733 GGAGGTTCATCATGTGCGCGGAATAAGAAAGGCGCTGGACTCTGCTGGTGTGACTTCC 792
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 219 GlyGlyProLeuValCysProAsnGlyAspGlyThrTyrSerLeuAlaGlyIleValSer 238
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 793 TGGGGTTTGGCTGTGTCGAGGCTGGAGAAACAATGTGAGGAAAAAGTATCAAGGATCC 852
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 239 TrpGlyIleGlyCysAlaGlnPro---ArgAsn----- 248
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 853 CTGGGATCTTCACAGACATATTAGTAAGTCTTCTTCTGGATC 894
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 249 ProGlyValTyrThrGlnValSerLysPheLeuAspTrpIle 262
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 5
O9VEY6 ID O9VEY6 PRELIMINARY; PRT; 787 AA.
AC O9VEY6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE SB gene product.
GN SB OR CG4316.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.

```





QY 157 TTCAAGTCATCTTGGAGAGCAAGTCAGTGGAGAGGTTCTCTATCCCTGGCAGGTATCT 216  
 DB 76 PheMetLysIleMetGlyGlyValAspAlaGluGlyLysTrpProTrpGlnValSer 95  
 QY 217 CTGAACAAGAGGAGCAAGCATATTTGTGGAGGAGCATCTCTACACACAGTGGGTATC 276  
 DB 96 ValArgValArgHisMetHisValCysGlySerLeuIleAsnSerGlnTrpValLeu 115  
 QY 277 ACCGGCGCTCAGTCATCTGAACACAGCAACATTTGTCTCTACTTTGAATGTACTGCTGGA 336  
 DB 116 ThrAlaAlaHisCysIleTyrSerArg-----IleGlnTyrAsnValLysValGly 132  
 QY 337 GAGTATGACTTAAGCCAGCAGACAGCCAGGAGCAAACTCTCACTATTGAAGTGTGATC 396  
 DB 133 AspArgSerValTyrArgGlnAsnThr---SerLeuValIleProIleLysThrIlePhe 151  
 QY 397 ATACATCCACATTTCTCCACCAAGAACCAATGAGTATGATATGTCCTTTGAAGATG 456  
 DB 152 ValHisProLysPheSerThrThrIleValValLysAsnAspIleAlaLeuLeuLysLeu 171  
 QY 457 GCTGGAGCCTTCCAAATTTGCCACTTTGTGGGCCCATATGCTCTCCAGAGCTCGGGGAG 516  
 DB 172 GlnHisProValAsnPheThrThrAsnIleTyrProValCysIleProSerGluSerPhe 191  
 QY 517 CAATTTGAGGCTGGTTTATTGTGTACAACTGCAGGCTGGGCCGCTTTAACTGAAGGTGGC 576  
 DB 192 ProValLysAlaGlyThrLysCysTrpValThrGlyTrpGlyLysLeuValProGlyAla 211  
 QY 577 -----GTCCTCTCACAAAGTCTTGCAGGAAGTGAATCTGCCTATTGACCTGGGAAGAG 630  
 DB 212 ProAspValProThrGluIleLeuGlnGluValAspGlnAsnValIleLeuTyrGluGlu 231  
 QY 631 TGTGTGCAGCTCTGTTAACTAAAGAGGCCATCAGTGGG-----AAG 675  
 DB 232 CysAsnGlu-----MetLeuLysLysAlaThrSerSerValAspLeuValLys 248  
 QY 676 ACCTTTCTTCCACAGGTTTCTCTGAGGAGGAGAGCAGCATGTCTCAGGAGATTCAGGA 735  
 DB 249 ArgGlyMetValCysGlyTyrLysGluArgGlyLysAspAlaCysGlnGlyAspSerGly 268  
 QY 736 GGTCACTCATGTC-----CGGATAAGAAAGGGCCCTGGAGCTCGGCTGGTGTGACT 789  
 DB 269 GlyProMetSerCysGluPheGluAsnLys-----TrpValGlnValGlyValVal 285  
 QY 790 TCCCTGGGTTTGGGCTGTGTGAGGCTGGAGAAACAATGTGAGAAAGTGTATCAAGGA 849  
 DB 286 SerTrpGlyIleSerCysGlyArg-----LysGly 295  
 QY 850 TCCCTGGGATCTTCAACAGCATATTAGTAAAGTGTCTTCCCTGGATC 894  
 DB 296 TyrProGlyValTyrThrAspValAlaPheTyrSerLysTrpLeu 310

## RESULT 8

Q9DAT3 ID Q9DAT3 PRELIMINARY: PRT; 624 AA.  
 AC Q9DAT3;  
 DT 01-JUN-2001 (T-EMBLrel. 17, Created)  
 DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)  
 DE 1600027G01Rik protein (RIKEN cDNA 1600027G01 gene).  
 GN 1600027G01Rik.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=PLACENTA;  
 RX MEDLINE=2108560; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Sakurai L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Guncicich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki K., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LIVER;  
 RA Strausberg R.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY.  
 CC EMBL; AK005546; BAB24114.1; -;  
 DR EMBL; BC019485; AAH19485.1; -;  
 DR HSSP; P00750; IRTF.  
 DR MEROPS; S01.213; -;  
 DR MGD; MGI:1919281; 1600027G01Rik.  
 DR InterPro; IPR000177; Apple.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR InterPro; IPR003014; PAN.  
 DR InterPro; IPR001254; Ser--protease\_Try.  
 DR Pfam; PF000024; PAN; 4.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00005; APPLIEDOMAIN.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00223; APPLE; 4.  
 DR SMART; SM00020; Tryp\_Spc; 1.  
 DR PROSITE; PS00495; APPLE; UNKNOWN\_3.  
 DR PROSITE; PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS; UNKNOWN\_1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 KW Hydrolase; Serine protease.  
 SQ SEQUENCE 624 AA; 69788 MW; 0BEDDEBC56009E97 CRC64;  
 Alignment Scores:  
 Pred. No.: 3.65e-40 Length: 624  
 Score: 473.50 Matches: 95  
 Percent Similarity: 53.82% Conservative: 46  
 Best Local Similarity: 36.26% Mismatches: 98  
 Query Match: 28.39% Indels: 23  
 DB: 11 Gaps: 6  
 US-09-735-713a-1 (1-921) x Q9DAT3 (1-624)  
 QY 127 AAGTACAGCTTGGAAATTTTAAACATTTTCACTGCGCATTTCTTGGAGGAGCAAGTG 186  
 DB 385 LysIleAsnPro-----ArgValValGlyGlyAlaAlaSer 396  
 QY 187 GAGAAGGGTTCCTATCCCTGGCAGGTATCTCTGAACAAAGGAGGAGCATATTGTGGGA 246  
 DB 397 ValHisGlyGluTrpProTrpGlnValThrLeuHisIleSerGlnGlyHisLeuLysGly 416  
 QY 247 GGAAGCATCTCTCACCACAGTGGTATCATCAGCGGCTCTACTGTCATTTGAAACAGAAC 306  
 DB 417 GlySerIleIleGlyAsnGlnTrpIleLeuThrAlaAlaHisCysPheSerGlyIleGlu 436  
 QY 307 ATTGTGCTACTTTGAATGTACTGCTGGAGAGTATGACCTTAAGCCAGACAGCCAGGA 366  
 DB 437 ThrProLysLysLeuArgValTyrGlyGlyIleValAsnGlnSerGluIleAsnGly 456  
 QY 367 GAGCAAACTCTCACTATTGAAGAACTGTCTATCATATACATATTTCTCCCAAGAAACCA 426

```

Db 457 ThrAlaPhePheArgValGlnGluMetIleIleHisAspGlnTyrThrAlaGlu--- 475
QY 427 ATGGACTATGATATGCGCTTTTGAAGATGCTGGAGCGCTTCCAAATTTGGCCACTTTGTG 486
Db 476 SerGlyTyrAspIleAlaLeuLeuLysLeuGluSerAlaMetAsnTyrThrAspPheGln 495
QY 487 GGGCCATATCTCTTCAGAGCTCGGGAGCAATTTGAGCGTGGCTTTATTTGACAACT 546
Db 496 ArgProIleCysLeuProSerLysGlyAspArgAsnAlaValHisThrGluCysTrpVal 515
QY 547 GCAGGCTGGGCCCTTAACTGAAGGTGGCTCTCTCACAAAGCTTTCAGGAAGTGAAT 606
Db 516 ThrGlyTrpGlyTyrThrAlaLeuArgGlyGluValGlnSerThrLeuGlnLysAlaLys 535
QY 607 CTGCCTATTTTGACCTGGGAGAGCTGTGGCAGCTCTGTTAACACTAAAGAGCCCATC 666
Db 536 ValProLeuValSerAsnGluGluCysGlnThrArgTyr-----ArgAlaHisLysIle 553
QY 667 AGTGGGAAGACCTTTCTTCACAGGTTTTCCTGATGGAGGAGAGAGCGCATGTCAGGGA 726
Db 554 ThrAsnLys---MetIleCysAlaGlyTyrLysGluGlyLysAspThrCysLysGly 572
QY 727 GATTCAGAGGTTCACTCATGTCCCGAATAAGAAAGGGCGCTGCTGCTGGTGGTGTG 786
Db 573 AspSerGlyGlyProLeuSerCysLys---TyrAsnGlyValTrpHisLeuValGlyIle 591
QY 787 ACTTCCTGGGTTTGGCTGTGCTGAGGCTGGAGGCTGGAGAAACAATGTGAGAAAGTATCAA 846
Db 592 ThrSerTrpGlyGluGlyCysGlyGlnLysGluArg----- 603
QY 847 GGATCCCTGGGATTCACAGACATAGTAAAGTGTCTTCCTGGATCCACAGAACATC 906
Db 604 -----ProGlyValTyrThrAsnValAlaLysTyrValAspTrpIleLeuGluLysThr 621
QY 907 CAAACT 912
Db 622 GlnThr 623

RESULT 9
Q91V47 PRELIMINARY; PRT; 624 AA.
AC Q91V47;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Coagulation factor XI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=B10.WR; TISSUE=LIVER;
RC MEDLINE=97385041; PubMed=9242536;
RX Gallani D., Sun M.F., Sun Y.;
RA "A comparison of murine and human factor XI.";
RL Blood 90:1055-1064(1997).
RN [2]
SEQUENCE FROM N.A.
RP STRAIN=B10.WR; TISSUE=LIVER;
RC Gallani D., Sun M.F., Sun Y.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF356627; AAK40233.1;
DR MEROPS; S01.213;
DR InterPro; IPR000177; Apple.
DR InterPro; IPR003014; PAN.
DR Pfam; PF00024; PAN; 4.
DR Pfam; PF00089; trypsin; 1.
DR PROSITE; PS00495; APPLE; UNKNOWN_2.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.

```

```

DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
SQ Hydrolyase; Serine protease.
KW SEQUENCE 624 AA; 69874 MW; 49D281BFAEC12A03 CRC64;

Alignment Scores:
Pred. No.: 1,53e-39 Length: 624
Score: 467.50 Matches: 95
Percent Similarity: 53.44% Conservative: 45
Best Local Similarity: 36.26% Mismatches: 99
Query Match: 28.03% Indels: 23
DB: 11 Gaps: 6

US-09-735-713a-1 (1-921) x Q91V47 (1-624)
QY 127 AAGGTACAGCCTTGGGAATTATTTTAAACATTTTCAGTCGCATTCTTGAGAGAACCAAGTG 186
Db 385 LysIleasnPro-----ArgValValGlyGlyAlaAlaSer 396
QY 187 GAGAAGGTTCCCTATCCCTGGCAGGTATCTCTGAAACAAAGCGAGAGCATATTTGTGGA 246
Db 397 ValHisGlyGluTrpProTrpGlnValThrLeuHisIleSerGlnGlyHisLeuCysGly 416
QY 247 GGAAGCATCTCTCACCACAGTGGGTGATCATCGCGGCTCCTGCTGCAATTCGCAAAACAGAAAC 306
Db 417 GlySerIleIleGlyAsnGlnTrpIleLeuThrAlaAlaHisCysPheSerGlyIleGlu 436
QY 307 ATTGTCTACTTGTGAATGTTACTGCTGGAGAGTATGACTTAAAGCCAGACAGACCCAGGA 366
Db 437 ThrProLysLysLeuArgValTyrGlyGlyIleValAsnGlnSerGluIleAsnGluGly 456
QY 367 GAGCAACTCTCACTATTGAACTGCATCATCATCATCATCTCCACCAAGAAACCA 426
Db 457 ThrAlaPhePheArgGluGlnGluMetIleIleHisAspGlnTyrThrAlaGlu--- 475
QY 427 ATGGACTATGATATGCGCTTTTGAAGATGCTGGAGCGCTTCCAAATTTGGCCACTTTGTG 486
Db 476 SerGlyTyrAspIleAlaLeuLeuLysLeuGluSerAlaMetAsnTyrThrAspPheGln 495
QY 487 GGGCCCATATGCTTCACAGAGCTGCGGAGCAATTTGAGGCTGGTGTATTTGTGTAACACT 546
Db 496 ArgProIleCysLeuProSerLysGlyAspArgAsnAlaValHisThrGluCysTrpVal 515
QY 547 GCAGGCTGGGCGCTTAACTGAAGTGGCTGCTCTCACAAAGTCTGCGAGAACTGAAT 606
Db 516 ThrGlyTrpGlyTyrThrAlaLeuArgGlyGluValGlnSerThrLeuGlnLysAlaLys 535
QY 607 CTGCCTATTTGACCTGGGAAGAGTGTGGCAGCTCTGTTAACACTAAAGAGGCCCATC 666
Db 536 ValProLeuValSerAsnGluGluCysGlnThrArgTyr-----ArgAlaHisLysIle 553
QY 667 AGTGGGAAGACCTTTCTTCACAGGTTTTCCTGATGGAGGAGAGACCATGTCAGGGA 726
Db 554 ThrAsnLys---MetIleCysAlaGlyTyrLysGluGlyGlyLysAspThrCysLysGly 572
QY 727 GATTCAGAGGTTCACTCATGTGGCGGAATAAGAAAGGGCGCTGCTGCTGGTGGTGTG 786
Db 573 AspSerGlyGlyProLeuSerCysLys---TyrAsnGlyValTrpHisLeuValGlyIle 591
QY 787 ACTTCCTGGGTTTGGGCTGTGCTGAGGCTGGAGCAATTTGAGGCTGGTGTATTTGTGTAACAA 846
Db 592 ThrSerTrpGlyGluGlyCysGlyGlnLysGluArg----- 603
QY 847 GGATCCCTGGGATTCACAGACATAGTAAAGTGTCTTCCTGGATCCACAGAACATC 906
Db 604 -----ProGlyValTyrThrAsnValAlaLysTyrValAspTrpIleLeuGluLysThr 621
QY 907 CAAACT 912
Db 622 GlnThr 623

RESULT 10
Q967X8 PRELIMINARY; PRT; 467 AA.
ID Q967X8

```

Db	364	GlySerMetSerValThrLeuGlnGluValAspValProValLeuThrAlaAlaCys
Qy	634	GTGCGACCTGTAAACACTAAAGAGGCCCATCAGTGGGAGACCTTTCTTTGGCACAGGT
Db	384	SerSerTrpTyrSerSerLeuThr-----AlaAsnMetMetCysAlaGly
Qy	694	TTTCCTCATGGAGGAGAGCGCATGTCAGGGAGATTTCAGAGGTTCTCACTCATGTGCCGG
Db	399	PheSerAsnGluGlyLysAspSerCysGlnGlyAspSerGlyGlyPro--MetValTyr
Qy	754	AATAAGAAAGGGCGCTGGACTCGCTGGCTGGTCTGACTCCCTGGGGTTGGCTGGGTCGA
Db	418	SerAlaThrSerAsnTyrGluGlnIleGlyValValSerTrpGlyArgGlyCysAlaArg
Qy	814	GGCTGGAGAAACAATGTGAGGAAAAGTGATCAAGGATCCCTCGGATCTTCACAGACATT
Db	438	-----ProglyPheProGlyValTyrAlaArgVal
Qy	874	AGTAAAGTCTTCTCCTGATCCAGCAACACATCCAAACTGGTAAAC
Db	448	ThrGluTyrLeuGluTyrIleAlaAlaAsn-----ThrGlyAsn

RESULT 11

Q8VDE0

ID	Q8VDE0	PRELIMINARY;	PRT;	453 AA.
AC	Q8VDE0			
DT	01-MAR-2002	(TtEMBLrel. 20, Created)		
DT	01-MAR-2002	(TtEMBLrel. 20, Last sequence update)		
DT	01-JUN-2002	(TtEMBLrel. 21, Last annotation update)		
DE	TPRSS3 protein.			
GN	TPRSS3.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RA	Guipponi M., Scamuffa N., Scott H.S., Rossier C., Antonarakis S.E.;			
RT	"Isolation and characterization of the mouse Tmprss3 gene."			
RL	Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AJ300738; CAC83350.1; -			
DR	MED; MGI:2155445; Tmprss3.			
DR	InterPro: IPR001314; Chymotrypsin.			
DR	InterPro: IPR002172; LDL_recept_A.			
DR	InterPro: IPR001254; Ser.protease_Try.			
DR	InterPro: IPR001190; Srcr_receptor.			
DR	Pfam: PF00057; ldl_recept_a; 1.			
DR	Pfam: PF00089; trypsin; 1.			
DR	PRINTS; PR00722; CHYMOTRYPSIN.			
DR	SMART; SM00192; LDLa; 1.			
DR	SMART; SM00202; SR; 1.			
DR	SMART; SM00202; Tryp_Spc; 1.			
DR	PROSITE; PS01209; LDLRA_1; UNKNOWN_1.			
DR	PROSITE; PS50068; LDLRA_2; 1.			
DR	PROSITE; PS50287; SRCR_2; 1.			
DR	PROSITE; PS50240; TRYPSIN_DOM; 1.			
DR	PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.			
DR	PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.			
SQ	SEQUENCE 453 AA; 49529 MW; 21E5697DC8781BD3 CRC64;			

Alignment Scores:

Pred. No.:	5,25e-39	Length:	453
Score:	462.00	Matches:	102
Percent Similarity:	51.62%	Conservative:	57
Best Local Similarity:	33.12%	Mismatches:	115
Query Match:	27.70%	Indels:	34
DB:	11	Gaps:	8

US-09-735-713A-1 (1-921) x Q8VDE0 (1-453)

Qy	1	ATGAGTCTCAAAATGCTTATAGCAGGACAAAGCTGATTTTACTAGGAATAGTCTTT	60
Db	170	ValSerIleAsnHisLeuLeuSerAspAspValThrAlaLeuHisHisSerValTyr	180

```

QY 61 TTTGACRAGGT---AAATCTGCARCTCTTTTCGCTCCCAAGCTCCAGCTGTGGGCAG 117
Db 190 MetargGluGlyCysThrSerGlyHisValValThrLeuLysCysSerAlaCysGlyThr 209
QY 118 AGTCTGTTAAGGTACAGCCTTGGAAATATTTTAAACATTTTCAGTCGCATTTCTTGGAGGA 177
Db 210 ArgThr-----GlyTyrSerProArgIleValGlyGly 220
QY 178 AGCAAGTGGAGAGGTTCTCTATCCCTGGCAGGTATCTCTGAACAAAGCCAGACAT 237
Db 221 AsnMetSerSerLeuThrGlnTrpProTrpGlnValSerLeuGlnPheGlnGlyTyrHis 240
QY 238 ATTTGTGGAGAAAGTCGTCACACAGCTGGGTATCATCAGCGCGCTCATCTCATGTGA 297
Db 241 LeuCysGlyGlySerIleIleThrProLeuTrpIleValThrAlaAlaHisCysValTyr 260
QY 298 AACAGAAACATTTGCTCTCTGAATTTACTGCTGGAGAGTATGACTTAAAGCCAGACA 357
Db 261 AspLeuTyrHisProLysSerTrpThrValGlnValGlyLeu-----ValSerLeuMet 278
QY 358 GACCCAGGAGAGCAAACTCTCACTATTGAACTGTCTATCATCATACATCCACATTTCTCCACC 417
Db 279 AspSerProValProSerHisLeuValGluLysIleIleTyrHisSerLysTyrLysPro 298
QY 418 AAGAAACCAATGACTATGATATTTGCCCTTTTGAAGATGGCTGGAGCCTTCCAAATTTGGC 477
Db 299 LysArg---LeuGlyAsnAspIleAlaLeuMetLysLeuSerGluProLeuThrPheAsp 317
QY 478 CACTTTGTGGGCCCATATGCTTCCAGAGCTGCGGGAGCAATTTGAGCTGGTTTTATT 537
Db 318 GluThrIleGlnProIleCysLeuProAsnSerGluLysAsnPheProAspGlyLysLeu 337
QY 538 TGTACAACTCGAGCTGGGGCCCTTAATCTGAAGTGGCGTCTCTCAACAAGCTTTGGCAG 597
Db 338 CysTrpThrSerGlyTrpGlyAlaThrGluAspGlyGlyAspAlaSerProValLeuAsn 357
QY 598 GAAGTGAATCTGCCTATTTTGACCTGGGAGAGTGTGGCAGCTCTGTAAACACTAAAG 657
Db 358 HisAlaAlaValProLeuIleSerAsnLysIleCys-----AsnHis 371
QY 658 AGCCCATCAGTGGG-----AAGACCTTTCTTTTGGACAGGTTTCTCTGATGGA 705
Db 372 ArgAspValTyrGlyGlyIleIleSerProSerMetLeuCysAlaGlyTyrLeuLysGly 391
QY 706 GGGAGAGAGCATGTCAGGAGATTCAGGAGTTCACATCATCTGCGGGAATAAGAAAGG 765
Db 392 GlyValAspSerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGluArg--- 410
QY 766 GCCTGGACTCTGGCTGGTGTACTTCCCTGGGCTTTGGGCTGTGGTGTGGAGGCTGGAGAAAC 825
Db 411 LeuTrpLysLeuValGlyAlaThrSerPheGlyIleGlyCys----- 424
QY 826 AATGTGAGGAAAGTATCAAGGATCCCTGGATCTTCACAGACATTAAGTAAAGTGCTT 885
Db 425 -----AlaGluValAsnLysProGlyValTyrThrArgIleThrSerPheLeu 440
QY 886 TCCTGGATCCAGCAACATCAAA 909
Db 441 AspTrpIleHisGluGlnLeuGlu 448
RESULT 12
Q8R0P5 PRELIMINARY; PRT; 638 AA.
AC Q8R0P5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_taxid=10090;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC026555; AAH26555.1; -.
SQ SEQUENCE 638 AA; 71382 MW; CC27C93F4B57C599 CRC64;

Alignment Scores:
Pred. No.: 1.49e-38 Length: 638
Score: 458.00 Matches: 94
Percent Similarity: 53.10% Conservative: 43
Best Local Similarity: 36.43% Mismatches: 103
Query Match: 27.46% Indels: 18
DB: 11 Gaps: 6

US-09-735-713a-1 (1-921) x Q8R0P5 (1-638)
QY 154 ATTTTCAGTCGCATTTCTTGAGGAAGCCCAAGTGGAGAGGTTTCCTATCCCTGGCAGGA 213
Db 387 IleAsnAlaArgIleValGlyGlyThrAsnAlaSerLeuGlyGlyTrpProTrpGlnVal 406
QY 214 TCTCTGAAACAAAGG-----CAGAAGCATATTTGTGGAGGAAGCATCGTCTCACCA 264
Db 407 SerLeuGlnValLysLeuValSerGlnThrHisLeuCysGlyGlySerIleIleGlyArg 426
QY 265 CAGTGGGTGATCACGGCGCTCATCTGCATTTGCCAAACAGAAACATTTGTCTACTTTGAAT 324
Db 427 GlnTrpValLeuThrAlaAlaHisCysPheAspGlyIleProTyrProAspValTrpArg 446
QY 325 GTTACTGCTGGAGAGTATGACTTAAGCCAGACAGCCAGAGAGCAAACTCTCACTATT 384
Db 447 IleTyrGlyGlyIleLeuSerLeuSerGluIleThrLysGluThrProSerArgIle 466
QY 385 GAAACTGTCTATCATATCATCCACATTTCTCCACCAGAAACCAATGACTATGATATGCC 444
Db 467 LysGluLeuIleIleHisGlnGlyTyrLysValSerGluGly---AsnTyrAspIleAla 485
QY 445 CTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCACTTTGTGGGGCCCATATGTCTTCCA 504
Db 486 LeuIleLysLeuGlnThrProLeuAsnTyrThrGluPheGlnLysProIleCysLeuPro 505
QY 505 GAGCTCGGGAGCAATTTGAGGCTGGTATTATTGTACAACTGCAGCTGGGGCCGCTTTC 564
Db 506 SerLysAlaAspThrAsnThrIleTyrThrAsnCysTrpValThrGlyTyrGlyThr 525
QY 565 ACTGAAGTGGGCTCTCTCTCACAACTCTTGCAAGTGAATCGCTCATTTTGGACCTGG 624
Db 526 LysGluGlnGlyGluThrGlnAsnIleLeuGlnLysAlaThrIleProLeuValProAsn 545
QY 625 GAAGAGTGTGGCGAGCTCTGTAAACACTAAAGAGGCCCATCAGTGGGAAGACCTTTCTT 684
Db 546 GluGluCysGlnLys-----LysTyrArgAspTyrValIleAsnLysGlnMetIle 562
QY 685 TGCACAGCTTTCTTCGATGGAGGAGACAGCATCTCAGGAGATTCAGGAGGTTTCATCTC 744
Db 563 CysAlaGlyTyrLysGluGlyGlyThrAspAlaCysLysGlyAspSerGlyGlyProLeu 582
QY 745 ATGTGCCCGGAATAAGAAAGGCGCTCTGGCTGTGGCTGTGCTACCTTCTCGGCTTGGGC 804
Db 583 ValCysLysHis---SerGlyArgTrpGlnLeuValGlyIleThrSerTrpGlyGly 601
QY 805 TGTGGTTCGAGCTGGAGAAACAATGTGAGGAAAGTGTATGAGGATCCCTGGGATCTTC 864
Db 602 CysAlaArg-----LysaspGln-----ProGlyValTyr 611
QY 865 ACACACATTAGTAAGTCCTTCTCGTGGATCCACGACACATCCCAACACTGGTAAC 918
Db 612 ThrLysValSerGluTyrMetAspTrpIleLeuGluLysThrGlnSerSerAsp 629
RESULT 13
Q99L44
ID Q99L44 PRELIMINARY; PRT; 339 AA.

```

AC Q99L44;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Similar to protease, serine, 8 (Prostasin).  
 GN PRSS8.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=129/SVJ, AND SWISS; TISSUE=LUNG;  
 RA Vergnese G.M., Caughey G.H.;  
 RT "Molecular cloning and characterization of mouse prostasin, a type I  
 RT membrane-associated serine protease of the gamma-tryptase/prostasin  
 RT gene family.";  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RA Kitamura K., Takefumi N., Kimio T.;  
 RT "mouse serine protease.";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY.  
 CC EMBL; BC003851; AA03851.1; -;  
 DR EMBL; AF378086; AAL06320.1; -;  
 DR EMBL; AF378085; AAL06319.1; -;  
 DR EMBL; AB038244; BAB82496.1; -;  
 DR HSSP; P00734; IUVS.  
 DR MEROPS; S01.159; -;  
 DR MGD; MGI:1923810; Prss8.  
 DR InterPro; IPR001314; Chymotrypsin.  
 DR Pfam; PF00089; trypsin; 1.  
 DR PRINTS; PR00722; CHYMOTRYPSIN.  
 DR SMART; SM00020; TRYP-SPC; 1.  
 DR PROSITE; PS00240; TRYPsin\_DOM; 1.  
 DR PROSITE; PS00134; TRYPsin\_HIS; UNKNOWN\_1.  
 DR PROSITE; PS00135; TRYPsin\_SER; 1.  
 KW Hydrolase; Protease; Serine protease.  
 SQ SEQUENCE 339 AA; 36216 MW; BC2DE88BC057AF10 CRC64;

Alignment Scores:  
 Pred. No.: 8.59e-38 Length: 339  
 Score: 450.00 Matches: 100  
 Percent Similarity: 48.15% Conservative: 43  
 Best Local Similarity: 33.67% Mismatches: 120  
 Query Match: 26.98% Indels: 34  
 DB: 11 Gaps: 8

US-09-735-713a-1 (1-921) x Q99L44 (1-339)

QY 34 CTGATTTTACTAGTAGTCTTTTGTGAACRAGGTAAATCTGCARCTCTTTGCTC 93  
 Db 17 IleLeuLeuLeuLeuGlyLeuLeu-----GlnSerGlyIleArgAlaAsp 31  
 QY 94 CCCAAAGCTCCCAAGTTGTGGTCAGAGCTGCTGTTAAGGTACAGCTTGGAAATTTTAAAC 153  
 Db 32 GlyThrGluAlaSerCysGly-----AlaValIleGlnPro----- 43  
 QY 154 ATTTTCAGTCGATCTTGGAGGAGCCCAAGTCGAGAGGGTTCTTCCTCCGCGAGGTA 213  
 Db 44 -----ArgIleThrGlyGlySerAlaLysProGlyGlnTrpProTrpGlnVal 60  
 QY 214 TCTCTGAACAAAGGAGAGCATATTTTGGAGGAGCATCTCTCCACCAAGTGGGTG 273  
 Db 61 SerIleThrTyraSpGlyAsnHisValCysGlySerLeuValSerAsnLysTrpVal 80

QY 274 ATCAGCGCGCTCAGTGCATTCGAAACAGAAACATCTGTCTACTTTGAATGTACTGCT 333  
 Db 81 ValSerAlaAlaHisCysPheProArgGluHisSerArgGluAlaTyrrGluValLysLeu 100  
 QY 334 GGAGACTATGACTTAAGCCAGACAGACCCAGAGAGAACTCTCACTATTGAACATGTC 393  
 Db 101 GlyAlaHisGlnLeuAspSerTyrrSerAsnAspThrValValHisThrValAlaGlnIle 120  
 QY 394 ATCATATCCACATTTCTCCACCAAGAACCAATGCATATATATATGTCCTTCCAGAGTCCGG 453  
 Db 121 IleThrHisSerSerTyrr---ArgGluGluGlySerGlnGlyAspIleAlaLeuIleArg 139  
 QY 454 ATGCTGGAGCCCTTCCAAATTTGGCCCATTTGTGGGCGCCATATCTCTCCAGAGTCCGG 513  
 Db 140 LeuSerSerProValThrPheSerArgTyrrIleArgProIleCysLeuProAlaAlaAsn 159  
 QY 514 GACCAATTTGAGGCTGGTTTATTTGTACAACTGCAGGCTGGGCGCCCTTAACATGAAGGT 573  
 Db 160 AlaSerPheProAsnGlyLeuHisCysThrValThrGlyTrpGlyHisValAlaProSer 179  
 QY 574 GCGCTCTCTCA-----CAAGTCTTGCAGGAAGTGAATCTGCTATTTTGACCTGGAA 627  
 Db 180 ValSerLeuGlnThrProArgProLeuGlnGlnLeuValProLeuIleSerArgGlu 199  
 QY 628 GAGTGTGTGCGAGCTCTGTTAACACTAAAGAGCGCCATCAGTGGGAAGACC----- 678  
 Db 200 ThrCysSerCysLeuTyrrAsnIleAsnAlaValProGluGluProHisThrIleGlnGln 219  
 QY 679 ---TTTCTTGCACAGGTTTCTGATGGAGGAGAGACGATGTCTCAGGAGATTCAGGA 735  
 Db 220 AspMetLeuCysAlaGlyTyrrValLysGlyLysAspAlaCysGlnGlyAspSerGly 239  
 QY 736 GGTTCATCATGTCCCGAATAAGAAGGCGCTGCACTCTGCTGGTGTGACTTCTCTGG 795  
 Db 240 GlyProLeuSerCys---PrometGluGlyIleTrpTyrrLeuAlaGlyIleValSerTrp 258  
 QY 796 GGTTTGGGCTGTGCTGCGAGGCTGGAGAACAACTGTGAGGAAAAGTGTCAAGATCCCT 855  
 Db 259 GlyAspAlaCysGlyAlaProAsnArg-----Pro 268  
 QY 856 GGATCTTCCACAGACATTAAGTAAAGTCTTCTGATCCATCCACGACATC 906  
 Db 269 GlyValTyrrThrLeuThrSerThrTyrrAlaSerTrpIleHisVal 285

RESULT 14  
 O97506 PRELIMINARY; PRT; 643 AA.  
 ID O97506;  
 AC O97506;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
 DE Kallikrein.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=PLASMA;  
 RA Kimura A., Kimura T., Okimura H., Hamabata T., Ohnishi J.,  
 RA Moriyama A., Takahashi K., Takahashi T.;  
 RT "Identification of porcine follipain as plasma kallikrein, and its  
 RT possible involvement in the production of bradykinin within the  
 RT follicles of porcine ovaries.";  
 RL Mol. Reprod. Dev. 57:79-87(2000).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY.  
 DR EMBL; AB022425; BAA37147.1; -;  
 DR HSSP; P00766; ICHG.  
 DR MEROPS; S01.212; -;  
 DR InterPro; IPR000177; Apple.  
 DR InterPro; IPR001314; Chymotrypsin.

```
DR InterPro; IPR003014; PAN.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00024; PAN; 4.
DR Pfam; PF00089; trypsin; 1.
DR PRINTS; PR00005; APPLEDOMAIN.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00223; APPLE; 4.
DR SMART; SM00020; Tryp_SPC; 1.
DR PROSITE; PS00495; APPLE; 4.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 643 AA; 7227 MW; AFF2923E3C3CB80A CRC64;

Alignment Scores:
Pred. No.: 2,63e-37 Length: 643
Score: 446.00 Matches: 90
Percent Similarity: 54.15% Conservative: 47
Best Local Similarity: 35.57% Mismatches: 98
Query Match: 26.74% Indels: 18
DB: Gaps: 5

US-09-735-713A-1 (1-921) x 097506 (1-643)
Qy 160 AGTCGATCTTGGAGGAGCAAGTGGAGAGGGTTCTCTATCCCTGGCAGGTATCTGTG 219
Db. 397 ThrArgilleValGlyThrAspSerPheLeuGlyGluTrpProTrpGlnValSerLeu 416
Qy 220 AACAAAGG-----CAGAACATATTTGTGGAGGAGCATCGCTCACACAGTGG 270
Db 417 GlnAlaLysLeuArgAlaGlnAsnHisLeuGlyGlySerIleIleGlyHisGlnTrp 436
Qy 271 GTGATCAGCGCGCTCACTGATGCAATGCAACAGAACATTTGTCTACTTTGAATGTTACT 330
Db 437 ValLeuThrAlaAlaHisCysPheAspGlyLeuSerLeuProAspIleTrpArgIleTyr 456
Qy 331 GCTGAGAGTAGTACTTAAGCCAGACAGACCCAGGAGCAAACTCTCACTATTGAAACT 390
Db 457 GlyGlyIleLeuAsnIleSerGluIleThrLysGluThrProPheSerGlnValLysGlu 476
Qy 391 GTCATATACATCCACATTTCTCCACCAGAACCAATGAGTACTATGATATGGCCCTTTG 450
Db 477 IleIleIleHisGlnAsnTyrLysIleLeuGlu---SerGlyHisAspIleAlaLeuLeu 495
Qy 451 AAGATGGCTGGAGCCTTCCCAATTTGGCCACTTTGTGGGCGCCATATGCTTCCAGAGCTG 510
Db 496 LysLeuGluThrProLeuAsnTyrThrAspPheGlnLysProIleCysLeuProSerArg 515
Qy 511 CGGAGCAATTTGAGCGTGGTTTATTGTTACAACACTGCAGGCTGGGCGCGCTTAACGTAA 570
Db 516 AspAspThrAsnValValTyrThrAsnCysTrpValThrGlyTrpGlyPheThrGluGlu 535
Qy 571 GGTGGCTCCTCTCACAGTCTTCAGAGAGTGAATCTGCTATTTGACCTGGGAGAG 630
Db 536 LysGlyGluIleGlnAsnIleLeuGlnLysValAsnIleProLeuValSerAsnGluGlu 555
Qy 631 TGTGGCGAGCTCTGTTAAACACTAAAGAGGCCCATCAGTGGGAAGACCTTTCTTTGCACA 690
Db 556 CysGlnLysSerTyrArgAspHisLys-----IleSerLysGlnMetIleCysAla 572
Qy 691 GGTTCCTCTGAGGAGGAGAGCGCATGTCCAGGAGATTTCAGGAGGTTCACATCTGTC 750
Db 573 GlyTyrLysGluGlyGlyLysAspAlaCysLysGlyGluSerGlyGlyProLeuValCys 592
Qy 751 CGGAATAAGAAAGGGCGCTGACACTCTGGCTGTGCTGACTCTCTGGGTTGGGCTGTGGT 810
Db 593 Lys---TyrAsnGlyIleTrpHisLeuValGlyThrThrSerTrpGlyGluGlyCysAla 611
Qy 811 CGAGGCTGGAGAAAACATGTGAGGAAAAGTGATCAAGGATCCCTGGATCTTCACAGAC 870
Db 612 Arg-----ArgGluGlnProGlyValTyrThrLys 621
```

```
Qy 871 ATTAGTAAAGTGCTTCTCGATCCACGACACATCCAA 909
Db 622 ValIleGluTyrMetAspTrpIleLeuGluLysThrGln 634
RESULT 15
QY0YK1
ID Q90YK1 PRELIMINARY; PRT; 433 AA.
AC Q90YK1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Coagulation factor VII.
OS Brachydanio rerio (zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21353085; PubMed=11459993;
RA Sheehan J., Templer M., Gregory M., Hanumanthaiah R., Troyer D.,
RA Phan T., Thankavel B., Jagadeeswaran P.;
RT "Demonstration of the extrinsic coagulation pathway in teleostei:
RT Identification of zebrafish coagulation factor VII.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:8768-8773(2001).
DR EMBL; AY040345; AAK74192.1; -.
DR MROPS; S01.215; -.
DR InterPro; IPR00561; EGF-like.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001254; Ser_protease_Try.
DR InterPro; IPR000294; VitK_dep_GLA.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00594; gla; 1.
DR Pfam; PF00089; trypsin; 1.
DR SMART; SM00179; EGF_CA; 2.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_2.
DR PROSITE; PS00011; GLU_CARBOXYLATION; UNKNOWN_1.
DR PROSITE; PS00240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; UNKNOWN_1.
KW Hydrolase; Serine protease.
SQ SEQUENCE 433 AA; 48680 MW; CD9D1B179601BA4C CRC64;

Alignment Scores:
Pred. No.: 6,97e-37 Length: 433
Score: 441.50 Matches: 103
Percent Similarity: 51.56% Conservative: 46
Best Local Similarity: 35.64% Mismatches: 105
Query Match: 26.47% Indels: 35
DB: Gaps: 9

US-09-735-713A-1 (1-921) x Q90YK1 (1-433)
Qy 61 TTTGAACRAGTAATCTGCARCTTTTCGTCCTCCCAAAGTCCAGTTGTGGCAGAGT 120
Db 172 PheProCysGlyLysValProLeuLeuGlnAlaGlyLysAlaAasp----- 187
Qy 121 CTGTTAAGGTACAGCCTTGAATATTTTAACATTTTCAGTCCCATCTTTGGAGGAAGC 180
Db 188 -----HisGlnValAspLeuArgSerArgIleValGlyGlySer 200
Qy 181 CAAGTGGAGAGGGTTCTCTATCCCTGGCAGGTATCTCTGAACAAAGCAGACATATT 240
Db 201 GlnCysProLysGlyHisCysProTrpGlnValLeuLeuLysTyrGlyLysGlyPhe 220
Qy 241 TGTGGAGAGACATGCTCTCACCACAGTGGGTGATCATCGGCGGCTCACTGCTGCAAC 300
Db 221 CysGlyValIleTyrLysProThrTrpIleLeuThrAlaAlaHisCysLeuGlyLys 240
Qy 301 AGAACATTTGTCTACTTTGATGTTACTGCTGGAGAGTATGCTTAAGCCAGACAGAC 360
Db 301 Arg-----ArgGluGlnProGlyValTyrThrLys 621
```

Db 241 LeuLys---ValIysPheLeuArgIleValAlaGlyGluHisAspLeuGluValAspGlu 259  
QY 361 CCAGGAGAGCAAACTCTCACTATTGAAACTGTCACTACATCCACATTTCTCCACCAAG 420  
Db 260 GlyThrGluGlnLeuIleGlnValAspGlnMetPheThrHisProAlaTyr---ValSer 278  
QY 421 AAACCAATGGACTATGATATGCCCTTTTGAAGATGGCTGGAGCCCTTCCAATTTGCCAC 480  
Db 279 GluThrAlaAspSerAspIleAlaLeuLeuArgLeuArgThrProIleValTyrSerVal 298  
QY 481 TTTGTGGGCCCATATGCTTCCAGAGCTGCCGGAGCAATTTGAGCTGGTGTATTTTGT 540  
Db 299 TyrAlaValProValCysLeuPro---LeuArgGluMetAlaGluArgGluLeuTrpAla 317  
QY 541 -----ACAACCTGCAGCTGGCGCGCTTAACTGAAGGTGGCGCTCTCTCACAA 588  
Db 318 ValSerLysHisThrValSerGlyTrpGlyLysArgSerGluAspGlyProThrSerArg 337  
QY 589 GTCTTCAGGAAGTGAATCTGCTATTTTGACCTGGGAAGAGTGTGTG---GCAGCTCTG 645  
Db 338 LeuLeuArgArgLeuLeuValProArgIleArgThrGlnGluCysValGlnValSerAsn 357  
QY 646 TTAACACTAAAGAGGCCCATCAGTGGGAAGACCTTTCTTTCACAGCTTTTCTGTATGGA 705  
Db 358 LeuThrLeuThr-----SerAsnMetPheCysAlaGlyTyrIleGluGly 372  
QY 706 GGGAGAGACGATGTCAGGAGATTCAGGAGGTTCACATGTCGCGGAATAGAAAGG 765  
Db 373 ArgGlnAspSerCysLysAspSerGlyGlyProLeuValThrArgTyrArgAspThr 392  
QY 766 GCCTGGACTCTGGCTGGTGTGACTTCTCTGGGTTTGGGTTGGTGTGCGAGGCTGGAGAAC 825  
Db 393 AlaPhe---LeuLeuGlyIleValSerTrpGlyLysGlyCysAlaArg----- 407  
QY 826 AATGTCAGGAAGATGATCAAGGATCCCTGGGATCTTCACAGACATTAGTAAAGTCTT 885  
Db 408 -----ProGlySerTyrGlyIleTyrThrArgValSerAsnTyrLeu 421  
QY 886 TCCTGGATCCAGCAACACATCCAAACT 912  
Db 422 GlnTrpIleArgGlnThrThrAsnThr 430

Search completed: March 25, 2003, 07:04:11  
Job time : 97.5 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2003, 04:15:25 ; Search time 1854 Seconds  
(without alignments)  
8045.331 Million cell updates/sec

Title: US-09-735-713a-1  
Perfect score: 921  
Sequence: 1 atgagctcacaagtcttat.....acatccaaactggtactaa 921

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :			
EST : *			
1:	em_estba:	*	
2:	em_esthum:	*	
3:	em_estin:	*	
4:	em_estnu:	*	
5:	em_estov:	*	
6:	em_estpl:	*	
7:	em_estro:	*	
8:	em_htc:	*	
9:	gb_est1:	*	
10:	gb_est2:	*	
11:	gb_htc:	*	
12:	gb_est3:	*	
13:	gb_est4:	*	
14:	gb_est5:	*	
15:	em_estfun:	*	
16:	em_estom:	*	
17:	gb_gss:	*	
18:	em_gss_hum:	*	
19:	em_gss_inv:	*	
20:	em_gss_pln:	*	
21:	em_gss_vrt:	*	
22:	em_gss_fun:	*	
23:	em_gss_mam:	*	
24:	em_gss_mus:	*	
25:	em_gss_other:	*	
26:	em_gss_pro:	*	
27:	em_gss_rod:	*	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	479.8	52.1	670	10	BB625475
2	110.8	12.0	582	17	AZ319576
3	104.4	11.3	585	17	AZ320103
4	80.4	8.7	2246	11	AK005546
5	76.6	8.3	596	9	AL633968
6	76.4	8.3	675	13	BM104667

7	76.4	8.3	679	13	BM104492
8	74.8	8.1	627	13	BM104591
9	74.8	8.1	642	14	BM890103
10	74.8	8.1	642	14	BQ258513
11	74.8	8.1	645	13	BM183510
12	74.8	8.1	652	13	BM102995
13	74.8	8.1	653	13	BM157100
14	74.8	8.1	656	13	BM157053
15	74.8	8.1	657	14	BQ075190
16	74.8	8.1	658	14	BM860233
17	74.8	8.1	664	14	BM957767
18	74.8	8.1	668	14	BM861171
19	74.8	8.1	676	14	BM958270
20	74.8	8.1	682	14	BM777442
21	74.2	8.1	532	13	BM315903
22	74.2	8.1	540	14	BM860769
23	73.2	7.9	574	13	BM102953
24	73.2	7.9	681	14	BM957651
25	72.2	7.8	688	10	AW342995
26	72.2	7.8	689	13	BJ093907
27	71.6	7.8	592	13	BI984661
28	71.6	7.8	597	13	BI877311
29	71.6	7.8	673	10	AW154001
30	71.2	7.7	553	10	AW421919
31	71.2	7.7	553	10	AW153658
32	70.6	7.7	638	10	AW153658
33	70.6	7.7	723	10	AW174475
34	70.4	7.6	557	9	AL595526
35	70	7.6	573	13	BM102765
36	70	7.6	708	10	AW153684
37	67	7.3	654	14	BM860845
38	66.2	7.2	432	13	BJ068921
39	64.4	7.0	564	10	AW153439
40	63.4	6.9	536	9	AI722029
41	62.2	6.8	538	10	AW342956
42	62.2	6.8	619	13	BI875854
43	62.2	6.8	3030	11	AK004939
44	61.8	6.7	504	9	AI721818
45	61.6	6.7	513	10	AW422658

ALIGNMENTS

RESULT 1 BB625475 670 bp mRNA linear EST 26-OCT-2001  
LOCUS BB625475 RIKEN full-length enriched, adult male epididymis Mus  
DEFINITION musculus cDNA clone 9230106D23 5', mRNA sequence.

ACCESSION BB625475.1 GI:16463737  
VERSION BB625475  
KEYWORDS EST.  
SOURCE house mouse.

ORGANISM

REFERENCE

AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takanashi,F., Takeda,Y., Tanaka,F., Toya,T., Muramatsu,M. and Hayashizaki,Y.  
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

TITLE

JOURNAL

COMMENT

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216

Email: genome-res@gsc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/  
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh  
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
 Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura  
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and  
 Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara  
 Y. and Hayashizaki, Y.  
 Computer-based methods for the mouse full-length cDNA  
 encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa  
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and  
 Hayashizaki, Y.  
 Computational Analysis of Full-Length Mouse cDNAs Compared with  
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)  
 Please visit our web site (http://genome.gsc.riken.go.jp) for  
 further details.  
 e mouse tissues.

#### FEATURES source

Location/Qualifiers  
 1..670  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="9230106D23"  
 /clone\_lib="RIKEN full-length enriched, adult male  
 epidiymis"  
 /sex="male"  
 /tissue\_type="epidiymis"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"

/note="Site 1: SalI; Site 2: BamHI; cDNA library was  
 prepared and sequenced in Mouse Genome Encyclopedia  
 Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in  
 RIKEN. Division of Experimental Animal Research in Riken  
 contributed to prepare mouse tissues. 1st strand cDNA was  
 primed with a primer [5'  
 GAGAGAGAGAGATCCAGAGCTTTTTTTTTTTTNN 3'], cDNA was  
 prepared by using trehalose thermo-activated reverse  
 transcriptase and subsequently enriched for full-length by  
 cap-trapper. cDNA went through one round of normalization  
 to Rot = 10.0 and subtraction to Rot = 185.0. Second  
 strand cDNA was prepared with the primer adapter of  
 sequence [5' GAGAGAGAGATCTCGAGTTAATAATTAATCCCCCCCCC  
 3']. cDNA was cloned into the xhoI and BamHI sites.  
 Vector: a modified pBluescript KS(+) after bulk excision  
 from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end:  
 BamHI"

BASE COUNT 180 a 154 c 169 g 167 t

Query Match 52.18; Score 479.8; DB 10; Length 670;  
 Best Local Similarity 82.5%; Pred. No. 4.6e-137;  
 Matches 550; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 125 TTAAGGTACAGCTTGGAAATTTATTAACATTTTCAGTCGCATCTTGGAGGAAGCCAAAG 184  
 Db 2 TTAAGCCCCACCTTCAGAAATTTACTTTAGCCCTTTTCAGTCGCATCTTGGAGGAAGCCAAAG 61  
 QY 185 TGGAGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAACAAGAGCAGAGCATATTTGTC 244  
 Db 62 TGGAAAGAGGGTTCCTATCTCTTGGCAAGTGTCTCTGAAACAAGAGCAGAGCAGATCTGTG 121  
 QY 245 GAGGAGGATCGTCTCACCACAGTGGGTGATCAGCGGGCTCAGTCATGTCACAAACAGAA 304

Db 122 GAGGAGCATTATCTCTTCACAGTGGGTGCATCACAGCGCTCATTTGCTAGCTTAACAGAA 181  
 QY 305 ACATTGTCTTACTTGAATTTACTGCTGGAGAGTAGTACTTAAGCCAGACACACCCAG 364  
 Db 182 ACATTGATTAACATTTTGAATGTAATCTGCTGGAGAACATGACTTGAGCCAAAGCCAGCCAG 241  
 QY 365 GAGAGCAACCTCTCATTGAAACTGTGATCATACATATCCACATTTCTCCACCAAGAAAC 424  
 Db 242 GAGAGCAACACTTGGCCATTGAAACCATCATACACCCGCGCTTCTCCACCAAGAAAC 301  
 QY 425 CAATGGACTATGATATTTGCCCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGCCACTTG 484  
 Db 302 CTATGATCTATGATATAGCCCTTCTGAAGATGGCTGGAACCTTCCAATTTGGACAGTTG 361  
 QY 485 TGGGGCCATATGCTTTCCAGAGCTGGGGAGCAATTTGAGGCTGGTGTATTTGTACAA 544  
 Db 362 TGAGGCTGTGTGCTTTCCAGAGCCAGGAGAACACTTAATGCTGGATTTATTTGTACAA 421  
 QY 545 CTGAGGCTGGGGCGCTTAACTGAAGTGGCGTCTCTCACAAAGTCTTGCAGGAAGTGA 604  
 Db 422 CTGAGGCTGGGGTGGCTTGTCCGAAGTGGCAGGCTCCACAAAGTCTTGCAGCAGTAA 481  
 QY 605 ATCTGCTATTTTACCTGGGAAGAGTGTGTGGCAGCTCTTAACTAAAGAGGCCCA 664  
 Db 482 ACCTGCCAAATTTGACTCAGGAGGAATGTAGGCGGTTCTGTGACCTTAAAGAAATCCCA 541  
 QY 665 TCAGTGGGAAGACCTTTCTTTCACAGGTTTCTCTGATGGAGGAGAGACCATGTGAG 724  
 Db 542 TTACTGGGAAGACCTTTCTCTGCACAGGATCCCTGTATGGAGGAGAGATGCTTTGTCAAG 601  
 QY 725 GAGATTGAGGAGTTCATCTATGTCGCGGAATAAGAAAGGGCGCTGGACTCTGCTGTGTG 784  
 Db 602 GAGATTCAAGAGGGGCACTTATGTCACCAATAGAAAAGGGCCCTGGACTTTGGCTGGAG 661  
 QY 785 TGACTTC 791  
 Db 662 TGACTTC 668

#### RESULT 2

AZ319576/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AZ319576 682 bp DNA linear GSS 29-SEP-2000  
 1M0039E20F Mouse 10kb plasmid UUGCLM library Mus musculus genomic  
 clone UUGCLM0039E20 F, DNA sequence.  
 AZ319576  
 AZ319576.1 GI:10370488  
 GSS.  
 house mouse.  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 682)  
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,  
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly  
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A.  
 and Wright, D., Weiss, R.  
 Mouse whole genome scaffolding with paired end reads from 10kb  
 plasmid inserts  
 Unpublished (2000)  
 Contact: Robert B. Weiss  
 University of Utah Genome Center  
 University of Utah  
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
 84112, USA  
 Tel: 801 585 5606  
 Fax: 801 585 7177  
 Email: ddunn@genetics.utah.edu  
 Insert Length: 10000 Std Error: 0.00  
 Plate: 0039 row: E column: 20  
 Seq primer: CGTTGTAACAGCGGCCACT  
 Class: plasmid ends  
 High quality sequence stop: 682.



KEYWORDS	HTC; CAP trapper.	TITLE	Direct Submission
SOURCE	Mus musculus (strain:C57BL/6J) adult female placenta cDNA to mRNA, clone:1600027G01.	JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
ORGANISM	Mus musculus	COMMENT	Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1		cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer 15' GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence 5' GAGAGAGAGCGCCCAATTATTCGAGTTAATTAATTAATCCCCCCCCC 3'. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end:XhoI; 3' end:SstI. Host: SOLR.
FEATURES	3	FEATURES	Location/Qualifiers
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	source	1. .2246
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		/organism="Mus musculus"
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		/strain="C57BL/6J"
MEDLINE	20499374		/db_xref="FANTOM_DB:1600027G01"
PUBMED	11042159		/db_xref="MGD:MGI:1904755"
REFERENCE	4		/db_xref="taxon:10090"
AUTHORS	Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaide, I., Pesole, G., Quackenbush, J., Schriml, L.M., Staubli, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S. and Hayashizaki, Y.	CDS	/clone_start=1
TITLE	Functional annotation of a full-length mouse cDNA collection		/protein_id="BAB24114.1"
JOURNAL	Nature 409 (6821), 685-690 (2001)		/db_xref="GI:12838178"
MEDLINE	21085660		/db_xref="MGD:MGI:1919281"
PUBMED	11217851		/translation="MTSLHQVLYFFFAVSSECVTKVKFDISFGQDLSSTVTPSAT YCRLVCTHPHRCLLFTFAESSDDPTKWFACILKDSVTEILPMVMTGAISYSEKQ CPQOLSTCKDVVNLDMKGMVNSVSNARCEQRTDDAHCQFTVATGCTVGFPSVD HRKMCILKTRTGTPTTITKLNGVSGFLSKGLNLACIRIDIPNTVLADINISV VAPDAVCRICHTHPTCTLEFFFSQAMPKESORHLCLKLTSEGLPSPTIKSHALS GFSLQCRHSVPVFCPSYNDIDFLGELDIDVVGQETCCQCTNNARCCQFTTYP SHRLCNERNRGRCYLKLSSGSPRILHGRGIGSYLRLCKMDNVCTKINPRVVG GAASVHGEPWPQVTLHSQHLGSGSIIGNOWILTAHCFSGIETPKLRVGGIYVQ NECTEATFPRVOEMIHDQYTTAESGYDIALKLESAMNYTDFQRPICLPKSGDRNA VHTTECVGTGWYALTARGEVQSTLOKAKVPLVSNECOTVRRHKITNMKTCAGYKEGG KDTCKGDSGSPILSCKYNGVHLVITSWGEGCGQKRPVGYTNVAKYVDHLEKTQTV "
REFERENCE	5 (bases 1 to 2246)	BASE COUNT	666 a 509 c 536 g 535 t
AUTHORS	Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.	ORIGIN	
		Query Match	8.7%; Score 80.4; DB 11; Length 2246;
		Best Local Similarity	48.5%; Pred. No. 6.4e-13;
		Matches 321; Conservative	0; Mismatches 326; Indels 15; Gaps 3;
		QY	150 TAACATTTTCAGTCGCATTTCTTGAGGAGGACCAAGTGGAGAGGGTTCTTCTATCCCTGGCA 209
		Db	1245 TAAATCAACCCAGAGTGGTAGGAGGAGCTGCGTCTGTTCCAGGTGAGTGGCATGGCA 1304
		QY	210 GGTATCTCTGAACAAGAGGAGAGCATATTTGTGGAGAGAGCATCTCTCACCACAGTG 269
		Db	1305 GGTGACTCTGCATCAGCAGGAGGACACCTCTGTGTGGAGGCTCCATCATTTGGAACCAATG 1364
		QY	270 GGTGATCACCGCGCTCAGTCGATTGCAACACAGAACAAATTGTGTCTACTTTGAATGTAC 329
		Db	1365 GATACTGACAGCAGCTCATTTGTTCTCTGGGATAGAGACACCTAAAAAGGTGGGTGCTA 1424

QY	330	TGCTGGAGAGTATGACTTTAAGCCAGACAGACCCAGGAGAGCAAACTCTCACTATTGAAAC	389
Db	1425	CGTGGCATTGTAAATCAATCAGAATAAATGAAGGAGACTGCTTCTTTCAGGGTTCGAAGA	1484
QY	390	TGTCATCATACATCCACATTTCTCCACCAAGAAACCAATGACATATGATATTGCCCCTTTT	449
Db	1485	AATGATAATTATGATCAGTATACGACAGCAGA---AAGTGGGTATGATATTGGCCCTGTT	1541
QY	450	GAAGATGGCTGGAGCCCTTCCAAATTTGGCCACTTTGTGGGGCCCATATGTCTTCCAGAGCT	509
Db	1542	AAAATGGAATCAGCCATGAATTACACAGATTTTTCAGCGGCCAATATGCCTGCCTTCCAA	1601
QY	510	CGGGGAGCAATTTGAGGCTGGTTTTATTGTACAACTCGAGGCTGGGGCCGCTTAACTGA	569
Db	1602	AGGAGATAGAAAGCGAGTGCACACAGAATGCTGGGTGACTGGATGGGGGTACACAGCACT	1661
QY	570	AGTGGCGCTCTCTCAAGACTTGCAGGAAGTGAATCTGCCTATTTTGGACCTGGGAAGA	629
Db	1662	AAGAGTGAAGTACAAAGTACTTCTCAGAAAGCCAAAGGTTCCATTGGTGTCAAATGAAGA	1721
QY	630	GTGTGTGGCAGCTCTGTTAAACACTAAAGAGGCCCATCAGTGGGAGAGCACTTTCTTTGCAC	689
Db	1722	ATG-----TCAGACAAGATACAGAAGACACAAAATAACCAATAGATGATCTGTC	1772
QY	690	AGGTTTTCTGATGGAGGAGAGACGCATGTCAGGGAGATTCAGAGGTTTCACTCATGTG	749
Db	1773	AGGCTACAAGAGGAGGGAAGGATACGTGCAAGGGAGATTCTGGAGG---GCGCCCTGTC	1829
QY	750	CCGGAATACGAAGGGCCCTGGACTCTGGCTGGTGTACTCTCTGGGGTTTGGGCTGTGG	809
Db	1830	CTGCAATACATGGGCTCTGGCACTTGGTGGGCATCACAGCTGGGCTGAAGCTGTGG	1889
QY	810	TC 811	
Db	1890	TC 1891	

RESULT 5	596 bp	mRNA	linear	EST 02-NOV-2001		
AL633968	AL633968	XGC-gastrula	Silurana tropicalis	cdna clone TGas012h03 5';		
LOCUS	mRNA sequence.					
DEFINITION						
ACCESSION	AL633968					
VERSION	AL633968.1	GI:16603451				
KEYWORDS	EST.					
SOURCE	western clawed frog.					
ORGANISM	Silurana tropicalis					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae; Xenopodinae; Silurana.					
AUTHORS	1 (bases 1 to 596)					
TITLE	Huckle, E., Taylor, R., Ashurst, J. L., Zorn, A. M. and Rogers, J.					
JOURNAL	Sanger Xenopus tropicalis EST project 2001 (10_2001)					
COMMENT	Unpublished (2001)					
	Contact: Huckle E					
	Sanger Centre					
	Hinxton, Cambridgeshire, CB10 1SA, UK					
	Email: trop@sanger.ac.uk					
	Sanger Xenopus tropicalis EST project 2001					
	TROPICALIS_SEQUENCE_ID: TGas012h03.sp6					
	Sequencing primer: SP6					
	This sequence is from a Xenopus Gene Collection (XGC) library					
	constructed by Aaron M. Zorn.					
FEATURES	Location/Qualifiers					
source	1..596					
	/organism="Silurana tropicalis"					
	/db_xref="taxon:8364"					
	/clone="TGas012h03"					
	/clone_lib="XGC-gastrula"					
	/dev_stage="gastrula (stages 10.5-13 mixed)"					
	/lab_host="Escherichia coli XL1-blue"					
	/note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cdna					

BASE COUNT	149 a	156 c	146 g	145 t	
ORIGIN					
Query Match	8.3%; Score 76.6; DB 9; Length 596;				
Best Local Similarity	51.1%; Pred. No. 3.6e-12;				
Matches	206; Conservative 0; Mismatches 194; Indels 3; Gaps 1;				
QY	155	TTTTACATCGCATCTCTTTGGAGGAAGCCAAAGTGGAGAGAGGGTTCCTATCCCTGGCAGGTAT	214		
Db	166	TTTCCACGTAGGATGTTTGGAGGAACCCACACTAGGCAAGGAGCATGGCCTTTGGCAGATCA	225		
QY	215	CTCTGAAACAAGGCGACAAGCATATTTGTGAGGAAGCATCGTCTCACCACAGTGGGTGA	274		
Db	226	GCTTAGAGTTAATGGCAGGCATATTTGTGAGGGTCCATAGTCTCTGACCACTGGATTC	285		
QY	275	TCACGGCGGCTCACTGCATTGCAACACAGAAACATTTGTCTACTTTGAATGTTACTGCTG	334		
Db	286	TGACAGCGGCCACTGTCATTGAACACCCGACATCTGGTTATGGAGTCCGCTCTTG	345		
QY	335	GAGATATGACTTAAGCCAGACAGACCCAGGAGAGCAAACTCTCACATATTGAAACTGTCA	394		
Db	346	CGCGGTACCAACTGTATGTGAAGAACCACCATGAGATCACAAATCAAGGTTACCGCCATCT	405		
QY	395	TCATACATCCACATCTCTCCACCAAGAAACAAATGGACTATGATATGGCCCTTTTGAAGA	454		
Db	406	ACGTGAATCCCACTTTGATGGCCCGCGGAGTG---GAGATATTGCTTTATTTGAAGC	462		
QY	455	TGGCTGGAGCCTTCCAAATTTGGCCACTTTGTGGGGCCCATATGCTTCCAGAGCTGCGGG	514		
Db	463	TGCTAGCCCATTAATATACGAATACATCTCCCAATATGCATGCGACTGCTACGG	522		
QY	515	AGCAATTGAGGTGGTGTATTTGTACAACCTGCAGCGTGGGG	557		
Db	523	CCACATTTCCCGCTGGGACCAATGCTGGGTAAACAGGATGGGG	565		
RESULT 6					
BM104667					
LOCUS	fv49h11.y1 Sugano SJD adult male Danio rerio cDNA clone 5411972 5'				
DEFINITION	similar to SW:TMS2_HUMAN O15393 TRANSMEMBRANE PROTEASE, SERINE 2 ; mRNA sequence.				
ACCESSION	BM104667				
VERSION	1				
KEYWORDS	EST.				
SOURCE	BM104667.1 GI:17035737				
ORGANISM	zebrafish.				
REFERENCE	Danio rerio				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes ; Cyprinidae; Danio.				
	1 (bases 1 to 675)				
	Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.				
TITLE	WashU zebrafish EST Project 1998				
COMMENT	Unpublished (1998)				
	Other_ESTs: fv49h11.x1				
	Contact: Stephen L. Johnson				
	Washington University School of Medicine				
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA				
	Tel: 314 286 1800				
	Fax: 314 286 1810				
	Email: zbrafish@watson.wustl.edu				
	Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution information can be found through the I.M.A.G.E.				





Custom primers for sequencing: 5' end primer  
CTTCTGCTCTAAAGTGGC and 3' end primer  
CGACCTGCAGCTCGAGCACA."

BASE COUNT 169 a 139 c 155 g 179 t  
ORIGIN

Query Match 8.1%; Score 74.8; DB 14; Length 642;  
Best Local Similarity 51.2%; Pred. No. 1.4e-11;  
Matches 175; Conservative 0; Mismatches 167; Indels 0; Gaps 0;  
QY 160 AGTGGCATTCTTGAGGAGCCCAAGTGGAGAGGGTTCTTATCCCTGGCAGGATCTCTG 219  
Db 196 AGCCGGATTATTGTGGAGAGAGGCTGGGCACATTCGTGGCCATGGCAGGTTCTCTG 255  
QY 220 AAACAAAGGAGAGAGCATATTGTGGAGGAGCATCTCTCACCACAGTGGGTGATCAGC 279  
Db 256 CAGTACATGATGCGCAACCTGTGGTGGTGCATCTCGACACGCGGTGGTTATTACC 315  
QY 280 GCGGCTCACTGCAATGCAACAGAAACATTTGTCTACTTTTGAATGTTACTGCTGGAGAG 339  
Db 316 GCGGCTCACTGCTTTAAACGTTTATAAAAGCCATCTATGTGGAAGCAGTGGTTGGATTG 375  
QY 340 TATGACTTAAAGCAGACAGACAGCCAGGAGAGCAAACTCTCACTATTGAAACTGTCAATCA 399  
Db 376 CACAATTGGATATGCGGAATGATCTCTACGGGAATCCATACAGTGTCAAAAATCTTC 435  
QY 400 CATCCACATTTCTCCACCAAGAACCAATGAGTATGATATGCTTCCCTTTTGAAGATGGCT 459  
Db 436 TCTCACAAGAACTACACACAGACAGAAACAGAAATGATATGCTTCCCTGCTGAACATGCAG 495  
QY 460 GGAGCCTTCCAATTGGCCCTTTGTGGGCCCCATATGCTTT 501  
Db 496 AGTCCGCTGGTTCAGTAAATTTGTGAGGCCCTATTGGCGCTT 537

## RESULT 10

BQ258513

LOCUS

DEFINITION

BQ258513 642 bp mRNA linear EST 06-MAY-2002  
fz2f02.y1 Sugano SJD adult male Danio rerio cDNA clone 5914899 5'  
similar to SW:TMS2\_HUMAN O15393 TRANSMEMBRANE PROTEASE, SERINE 2 ;  
mRNA sequence.

BQ258513.1 GI:20459272

EST.

SOURCE

zebrafish.

ORGANISM

Danio rerio

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes

; Cyprinidae; Danio.

1 (bases 1 to 642)

Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy

, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood

, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B.,

Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E.,

Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.

and Willson, R.

WashU Zebrafish EST Project 1998

Unpublished (1998)

Contact: Stephen L. Johnson

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: zbrfish@wustl.edu

Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA

Sequencing by: Washington University Genome Sequencing Center Clome

Consortium/LLN, send email to: info@image.llnl.gov

Seq primer: T3 ET from Anehrsam

High quality sequence stop: 526.

Location/Qualifiers

1..642

/organism="Danio rerio"

source

/db\_xref="taxon:7955"  
/clone="5914899"  
/clone\_lib="Sugano SJD adult male"  
/sex="male"  
/tissue\_type="whole body"  
/dev\_stage="adult"  
/lab\_host="DH10B (phage resistant)"  
/note="Vector: pME18S-FL3; Site 1: DraIII (CACCATGTG);  
Site 2: DraIII (CACTGTGTG); 1st strand cDNA was primed  
with an oligo(dT) primer (ATGTGGCTTTTCTTTTCTTTT);  
double-stranded cDNA was ligated to a DraIII adaptor  
[TGTGGCTTACTGG], digested and cloned into distinct DraIII  
sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site  
CACCATGTG). XhoI should be used to isolate the cDNA  
insert. Size selection was performed to exclude fragments  
<1.5kb. Library constructed and donated by Dr. Sumio  
Sugano (University of Tokyo Institute of Medical Science).  
Custom primers for sequencing: 5' end primer  
CTTCTGCTCTAAAGTGGC and 3' end primer  
CGACCTGCAGCTCGAGCACA."

BASE COUNT 170 a 138 c 155 g 179 t  
ORIGIN

Query Match 8.1%; Score 74.8; DB 14; Length 642;  
Best Local Similarity 51.2%; Pred. No. 1.4e-11;  
Matches 175; Conservative 0; Mismatches 167; Indels 0; Gaps 0;  
QY 160 AGTGGCATTCTTGAGGAGCCCAAGTGGAGAGGGTTCTTATCCCTGGCAGGATCTCTG 219  
Db 199 AGCCGGATTATTGTGGAGAGAGGCTGGGCACATTCGTGGCCATGGCAGGTTCTCTG 258  
QY 220 AAACAAAGGAGAGAGCATATTGTGGAGGAGCATCTCTCACCACAGTGGGTGATCAGC 279  
Db 259 CAGTACATGATGCGCAACCTGTGGTGGTGCATCTCGACACGCGGTGGTTATTACC 318  
QY 280 GCGGCTCACTGCAATGCAACAGAAACATTTGTCTACTTTTGAATGTTACTGCTGGAGAG 339  
Db 319 GCGGCTCACTGCTTTAAACGTTTATAAAAGCCATCTATGTGGAAGCAGTGGTTGGATTG 378  
QY 340 TATGACTTAAAGCAGACAGACAGCCAGGAGAGCAAACTCTCACTATTGAAACTGTCAATCA 399  
Db 379 CACAATTGGATATGCGGAATGATCTCTACGGGAATCCATACAGTGTCAAAAATCTTC 438  
QY 400 CATCCACATTTCTCCACCAAGAAACCAATGGACTATGATATGCTTCCCTTTTGAAGATGGCT 459  
Db 439 TCTCACAAGAACTACAAACAGACAGAAACAGAGAAATGATATGCTTCCCTGCTGAACATGCAG 498  
QY 460 GGAGCCTTCCAATTGGCCCTTTGTGGGCCCCATATGCTTT 501  
Db 499 AGTCCGCTGGTTCAGTAAATTTGTGAGGCCCTATTGGCGCTT 540

## RESULT 11

BM183510

LOCUS

DEFINITION

BM183510 645 bp mRNA linear EST 11-DEC-2001  
fv64d10.y1 Sugano SJD adult male Danio rerio cDNA clone 5413291 5'  
similar to SW:TMS2\_HUMAN O15393 TRANSMEMBRANE PROTEASE, SERINE 2 ;  
mRNA sequence.

BM183510.1 GI:17514468

EST.

KEYWORDS

SOURCE

zebrafish.

ORGANISM

Danio rerio

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes

; Cyprinidae; Danio.

1 (bases 1 to 645)

Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy

, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood

, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B.,

Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E.,

Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.

and Willson, R.



```

Db 434 TCTCACAAGAACTACAACCAAGAACAAACGAGAAATGATATGGCTGCTGAAACTGCAG 493
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 460 GGAGCCTTCCAATTGGCCACTTTGGGGCCCATATGTCTT 501
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 494 AGTCGGCTGGTTTCAGTAAATTTGTGAGGCGCTATTGGCGTT 535

RESULT 13
BM157100
LOCUS BM157100 653 bp mRNA linear EST 03-DEC-2001
DEFINITION fV44c08-y1 Sugano SJD adult male Danio rerio cDNA clone 5411319 5'
similar to SW:TM52_HUMAN O15393 TRANSMEMBRANE PROTEASE, SERINE 2 ;
mRNA sequence.
ACCESSION BM157100
VERSION BM157100.1 GI:17241395
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
1 (bases 1 to 653)
Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
and Wilson,R.
WashU Zebrafish EST Project 1998
Unpublished (1998)
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrfish@wustl.edu
Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
Sequencing by: Washington University Genome Sequencing Center Clone
distribution information can be found through the I.M.A.G.E.
Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: T3 ET from Amersham
High quality sequence stop: 521.
Location/Qualifiers
1 .653
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="5411319"
/clone_lib="Sugano SJD adult male"
/sex="male"
/tissue_type="whole body"
/dev_stage="adult"
/lab_host="DH10B (phage resistant)"
/notes="vector: pME18S-FL3; Site.1: DraIII (CACCATGTG);
Site.2: DraIII (CACTGTGTG); 1st strand cDNA was primed
with an oligo(dT) primer [ATGCGCCTTTTTTTTTTTTTT];
double-stranded cDNA was ligated to a DraIII adaptor
[GTGTGCTTACTGCG], digested and cloned into distinct DraIII
sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site
CACCATGTG). XhoI should be used to isolate the cDNA
insert. Size selection was performed to exclude fragments
<1.5kb. Library constructed and donated by Dr. Sumio
Sugano (University of Tokyo Institute of Medical Science).
Custom primers for sequencing: 5' end primer
CTTCTGCTCTTAAAGCTGCG and 3' end primer
CGACCTGCAGCTCGACACA."
BASE COUNT 169 a 141 c 157 g 186 t
ORIGIN

Query Match 8.1%; Score 74.8; DB 13; Length 653;
Best Local Similarity 51.2%; Pred. No. 1.4e-11;
Matches 175; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

```

```

Qy 160 AGTCGCATTTCTGGAGGAAGCAAGTCGAGAGGGTTCCCTATCCCTGGCAGGTATCTCTG 219
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 202 AGCCGATTTATTTGGTGAAGAGGAGCTGGGCACATTCGTGGCGCATGGCAGGTTCTCTG 261
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 220 AACAAAGGAGAGACATATTTGTGGAGGAGACATCTCTCACCACACAGTGGGTGATCAG 279
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 262 CAGTACAAATGATGTGCCAACCTGTGGTGGTCCCATCTCGACACAGCGGTGGGTTATTACC 321
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 280 GCGCTCACTGCTATTGCAACACAGAACATTTGTGCTACTTTGAATGTTACTGCTGGAGAG 339
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 322 GCCGTCACCTGCTTTAAACGTTTATAAAAGCCATCTATGTGGAACAGCGTGGTGGATTG 381
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 340 TATGACTTAAGCCACAGACAGACCCAGGAGAGCAAACTCTCACTATTGAAACTGTCTACATA 399
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 382 CACAATTTGGATATGCGAATGAATCTCACGGGAATCCATACAGGTGCAAAAATCTTC 441
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 400 CATCCACATTTCTCCACCAAGAACCAATGAGTACTATATTTGCCCTTTTGAAGATGGCT 459
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 442 TCTCACAAGAACTACAACCAAGACAGACAAAGAGAAATGATATGGCTGCTGCTGAAACTGCAG 501
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 460 GGAGCCTTCCAATTTGGCCACTTTGTGGGCGCCCATATGTCTT 501
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 502 AGTCCGCTGGTTCAGTAAATTTGTGAGGCGCTATTGGCGTT 543
      || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 14
BM157053
LOCUS BM157053 656 bp mRNA linear EST 03-DEC-2001
DEFINITION fV43g06-y1 Sugano SJD adult male Danio rerio cDNA clone 5411506 5'
similar to SW:TM52_HUMAN O15393 TRANSMEMBRANE PROTEASE, SERINE 2 ;
mRNA sequence.
ACCESSION BM157053
VERSION BM157053.1 GI:17241305
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
1 (bases 1 to 656)
Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
and Wilson,R.
WashU Zebrafish EST Project 1998
Unpublished (1998)
Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrfish@wustl.edu
Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA
Sequencing by: Washington University Genome Sequencing Center Clone
distribution information can be found through the I.M.A.G.E.
Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: T3 ET from Amersham
High quality sequence stop: 524.
Location/Qualifiers
1 .656
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="5411506"
/clone_lib="Sugano SJD adult male"
/sex="male"
/tissue_type="whole body"
/dev_stage="adult"
/lab_host="DH10B (phage resistant)"
/notes="vector: pME18S-FL3; Site.1: DraIII (CACCATGTG);
Site.2: DraIII (CACTGTGTG); 1st strand cDNA was primed
with an oligo(dT) primer [ATGCGCCTTTTTTTTTTTTTT];
double-stranded cDNA was ligated to a DraIII adaptor
[GTGTGCTTACTGCG], digested and cloned into distinct DraIII
sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site
CACCATGTG). XhoI should be used to isolate the cDNA
insert. Size selection was performed to exclude fragments
<1.5kb. Library constructed and donated by Dr. Sumio
Sugano (University of Tokyo Institute of Medical Science).
Custom primers for sequencing: 5' end primer
CTTCTGCTCTTAAAGCTGCG and 3' end primer
CGACCTGCAGCTCGACACA."
BASE COUNT 169 a 141 c 157 g 186 t
ORIGIN

Query Match 8.1%; Score 74.8; DB 13; Length 653;
Best Local Similarity 51.2%; Pred. No. 1.4e-11;
Matches 175; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

```

double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed and donated by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAGCTGCG and 3' end primer CGACCTGCAGCTCGAGCACAC.

BASE COUNT	174 a	142 c	158 g	182 t
ORIGIN				
Query Match	8.1%;	Score 74.8;	DB 13;	Length 656;
Best Local Similarity	51.2%;	Pred. No. 1.4e-11;		
Matches 175;	Conservative 0;	Mismatches 167;	Indels 0;	Gaps 0;
QY 160	AGTCGCATTCTTGGAGGAGCCAAAGTGGAGAGGGTTCTCTATCCCTGGCAGGTATCTCTG	219		
Db 195	AGCCGGATTATTGGTGGAAAGGAGGCTGGGCACATTCGTGGCCATGGCAGGTTTCTCTG	254		
QY 220	AAACAAGGCAGAGCATATTTGTGGAGGAAGCATGCTCTCACCACAGTGGGTGATCAGC	279		
Db 255	CAGTACAATGATGTGCCAACCTGTGGTGGTGGCCATCTCGACAGCGGTGGGTATTTACC	314		
QY 280	GGGGCTCACTGCTTCAACACAGAAACATTTGTCTTACTTTGAATTTACTGTCTGCGAGAG	339		
Db 315	GCCGGTCACTGCTTTAAACGTTTATAAAGCCATCTATGTGACAGCATGGTGTGATTTG	374		
QY 340	TATGACTTAAGCCAGACAGACCCAGAGAGCAAACTCTCACTATTGAAACTGTCAATCATA	399		
Db 375	CACAATTTGGATAATTCGAATGAATCCTCACGGGAATCCATACAGTGCAGAAATCTTC	434		
QY 400	CATCCACATTTCTCCACCAAGAACCAATGGACTATGATATGCCCCTTTTGAAGTGGCT	459		
Db 435	TCTCAAGAAGCACTACACCAAGAGAGACAAAGAGATGATTTGCCCTGCTGAAACTGCAG	494		
QY 460	GGAGCCCTCCCAATTTGGCCACTTTTGGGGCCCATATGCTT	501		
Db 495	AGTCCCGCTGTTTTCAGTAATTTGTGAGCCCTATTGGCGTT	536		

RESULT 15	
LOCUS	BQ075190 657 bp mRNA linear EST 02-APR-2002
DEFINITION	fy43a08.y2 Sugano SJD adult male Danio rerio cDNA clone 5602455.5, similar to SWTMS2_HUMAN O15393 TRANSMEMBRANE PROTEASE, SERINE 2 ; mRNA sequence.
ACCESSION	BQ075190
VERSION	BQ075190.1 GI:19904227
KEYWORDS	EST.
SOURCE	zebrafish.
ORGANISM	Danio rerio
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes ; Cyprinidae; Danio.
AUTHORS	1 (bases 1 to 657) Clark, M., Johnson, S.L., Lehrach, H., Lee, R., Li, F., Marra, M., Eddy, S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.
TITLE	WashU zebrafish EST Project 1998
JOURNAL	Unpublished (1998)
COMMENT	Contact: Stephen L. Johnson Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: zbrafish@wustl.edu Library constructed by Dr. Sumio Sugano and Dr. Koichi Kawakami DNA Sequencing by: Washington University Genome Sequencing Center Clone

distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov  
Seq primer: T3 ET from Amersham  
High quality sequence stop: 522.

FEATURES

source  
1..657  
/organism="Danio rerio"  
/db\_xref="taxon:7955"  
/clone\_lib="Sugano SJD adult male"  
/sex="male"  
/tissue\_type="whole body"  
/dev\_stage="adult"  
/lab\_host="DH10B (phage resistant)"  
/note="Vector: pME18S-FL3; Site 1: DraIII (CACCATGTG); Site 2: DraIII (CACTGTGTG); 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCCTTTTTTTTTTTTTTT]; double-stranded cDNA was ligated to a DraIII adaptor [TGTGGCTACTGG], digested and cloned into distinct DraIII sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should be used to isolate the cDNA insert. Size selection was performed to exclude fragments <1.5kb. Library constructed and donated by Dr. Sumio Sugano (University of Tokyo Institute of Medical Science). Custom primers for sequencing: 5' end primer CTTCTGCTCTAAAGCTGCG and 3' end primer CGACCTGCAGCTCGAGCACAC."

BASE COUNT 171 a 141 c 159 g 186 t  
ORIGIN

Query Match	8.1%;	Score 74.8;	DB 14;	Length 657;
Best Local Similarity	51.2%;	Pred. No. 1.4e-11;		
Matches 175;	Conservative 0;	Mismatches 167;	Indels 0;	Gaps 0;
QY 160	AGTCGCATTCTTGGAGGAGCCAAAGTGGAGAGGGTTCTCTATCCCTGGCAGGTATCTCTG	219		
Db 202	AGCCGGATTATTGGTGGAAAGGAGGCTGGGCACATTCGTGGCCATGGCAGGTTTCTCTG	261		
QY 220	AAACAAGGCAGAGCATATTTGTGGAGGAAGCATGCTCTCACCACAGTGGGTGATCAGC	279		
Db 262	CAGTACAATGATGTGCCAACCTGTGGTGGTGGCCATCTCGACCGCGGTGGGTATTTACC	321		
QY 280	GGGGCTCACTGCTTCAACACAGAAACATTTGTCTTACTTTGAATTTACTGTCTGGAGAG	339		
Db 322	GCCGGTCACTGCTTTAAACGTTTATAAAGCCATCTATGTGGAACCGAGTGGATTG	381		
QY 340	TATGACTTAAGCCAGACAGACCCAGAGAGCAAACTCTCACTATTGAAACTGTCAATCATA	399		
Db 382	CACAATTTGGATAATTCGAATGAATCCTCACGGGAATCCATACAGGTGCAGAAATCTTC	441		
QY 400	CATCCACATTTCTCCACCAAGAACCAATGGACTATGATATGCCCTTTTGAAGTGGCT	459		
Db 442	TCTCAAGAAGCACTACACCAAGAGAGAAAGAGATGATTTGCCCTGCTGAAACTGCAG	501		
QY 460	GGAGCCCTCCCAATTTGGCCACTTTTGGGGCCCATATGCTT	501		
Db 502	AGTCCCGCTGTTTTCAGTAATTTGTGAGCCCTATTGGCGTT	543		

Search completed: March 25, 2003, 06:30:34  
Job time : 1860 secs

**THIS PAGE BLANK (USPTO)**

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	467	28.0	454	4	US-09-518-046-2	Sequence 2, Appli
2	460	27.6	248	4	US-08-944-483-63	Sequence 63, Appl
3	446.5	26.8	455	4	US-09-261-416-2	Sequence 2, Appli
4	446	26.7	638	2	US-08-681-151-3	Sequence 3, Appli
5	442.5	26.5	798	1	US-08-200-900A-2	Sequence 2, Appli
6	442.5	26.5	798	5	PCF-US94-00616-2	Sequence 2, Appli
7	440	26.4	416	2	US-09-000-846-2	Sequence 2, Appli
8	439.5	26.3	328	4	US-09-386-642-11	Sequence 11, Appl
9	437.5	26.2	256	2	US-09-027-337-3	Sequence 3, Appli
10	437.5	26.2	256	4	US-09-644-600-3	Sequence 3, Appli
11	436	26.1	319	4	US-09-386-642-12	Sequence 12, Appl
12	432.5	25.9	255	4	US-08-944-483-67	Sequence 67, Appl

```
Db 171 ValSerIleAspHisLeuLeuProAspAspLysValThrAlaLeuHisHisSerValTyr 190
Qy 61 TTTGAACRAGT---AAATCTGCARCTCTTTTCGCTCCCAAGCTCCAGTTGTGGCAG 117
Db 191 ValArgGluGlyCysAlaSerGlyHisValThrLeuGlnCysThrAlaCysGlyHis 210
Qy 118 AGTCTGTTAAAGTACAGCCCTTGAATATTTTAACTATTTTCAGTCGCTCTTTGGAGGA 177
Db 211 ArgArg-----GlyTyrSerSerArgIleValGlyGly 221
Qy 178 ACCCAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAGGCGAAGCAT 237
Db 222 AsnMetSerLeuLeuSerGlnTyrProTrpGlnAlaSerLeuGlnPheGlnGlyTyrHis 241
Qy 238 ATTTGTGGAGGAGCATGCTCACCAGTGCGGTGATCAGCGGCTCACTGCTATTGCA 297
Db 242 LeuGlyGlySerValIleThrProLeuTrpIleThrAlaAlaHisCysValTyr 261
Qy 298 AACAGAAACATTTGCTACTTTTGAATGTTACTGCTGGAGAGTATGACTTAAAGCCAGACA 357
Db 262 AspLeuTyrLeuProLysSerTrpThrIleGlnValGlyLeuValSerLeuLeuAsp--- 280
Qy 358 GACCCAGGAGACAACTCCTATTTGAACACTGATCATCATCATCATCATCATCATCAT 417
Db 281 AsnProAlaProSerHisLeu---ValGluLysIleValTyrHisSerLysTyrLysPro 299
Qy 418 AAGAAACCAATGGACTATGATATTCCTTTTGAAGATGGCTGCAGCTTCCAATTTGGC 477
Db 300 LysArg---LeuGlyAsnAspIleAlaLeuMetLysLeuAlaGlyProLeuThrPheAsn 318
Qy 478 CACTTTGTGGGGCCCATATGCTTCCAGAGCTGCGGGAGCAATTTAGGCTGTTTATT 537
Db 319 GluMetIleGlnProValCysLeuProAsnSerGluGluAsnPheProAspGlyLysVal 338
Qy 538 TGTACAACTGAGGCTGGGCGGCTTAACTGAAGTGGCTGCTCTCTCACAAGTCTTCGAG 597
Db 339 CysTrpThrSerGlyTrpGlyAlaThrGluAspGlyGlyAspAlaSerProValLeuAsn 358
Qy 598 GAAGTGAATCTGCCTATTTTGAACCTGGAGAGTGTGTGGCAGCTCTGTAAACACTAAAG 657
Db 359 HisAlaAlaValProLeuIleSerAsnLysIleCys-----AsnHis 372
Qy 658 AGGCCCATCACTGGG-----AAGACCTTTCTTGGCAGAGTTCCTGATGGA 705
Db 373 ArgAspValTyrGlyGlyIleIleSerProSerMetLeuCysAlaGlyTyrLeuThrGly 392
Qy 706 GGGAGAGACGATCTCAGGGAGATTACAGAGTTCACATGTCGCCGAATAAGAAAGG 765
Db 393 GlyValAspSerCysGlnGlyAspSerGlyGlyProLeuValCysGlnGluArg--- 411
Qy 766 GCCTGGACTCTGGCTGTGCTACTTCTCTGGGTTTGGGCTGTGCTGAGGCTGGAGAAAC 825
Db 412 LeuTrpLysLeuValGlyAlaThrSerPheGlyIleGlyCys----- 425
Qy 826 AATGTGAGAAAGTATCAAGGATCCCTGGGATCTTCACAGACATTAAGTAAAGTCTT 885
Db 426 -----AlaGluValAsnLysProGlyValTyrThrArgValThrSerPheLeu 441
Qy 886 TCCTGGATCCACAGACATCCAA 909
Db 442 AspTrpIleHisGluGlnMetGlu 449
```

## RESULT 2

```
US-08-944-483-63
; Sequence 63, Application US/08944483
; Patent No. 6232456
; GENERAL INFORMATION:
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
; APPLICANT: GRANADOS, EDWARD N.
; APPLICANT: KLASS, MICHAEL R.
; APPLICANT: RUSSELL, JOHN C.
```

```
; APPLICANT: STEWART, KENT D.
; APPLICANT: STROUPE, STEVEN D.
; TITLE OF INVENTION: NOVEL SERINE PROTEASE REAGENTS
; TITLE OF INVENTION: AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
; TITLE OF INVENTION: OF THE PROSTATE
; NUMBER OF SEQUENCES: 76
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Abbott Laboratories
; STREET: 100 Abbott Park Road
; CITY: Abbott Park
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,483
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Becker, Cheryl L.
; REGISTRATION NUMBER: 35,441
; REFERENCE/DOCKET NUMBER: 6183.US.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 847/935-1729
; TELEFAX: 847/938-2623
; TELEX:
; INFORMATION FOR SEQ ID NO: 63:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 248 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6232456e
; US-08-944-483-63
```

## Alignment Scores:

Pred. No.:	3,4e-42	Length:	248
Score:	460.00	Matches:	90
Percent Similarity:	55.12%	Conservative:	50
Best Local Similarity:	35.43%	Mismatches:	96
Query Match:	27.58%	Indels:	18
DB:	4	Gaps:	5

US-09-735-713A-1 (1-921) x US-08-944-483-63 (1-248)

```
Qy 166 ATCTTGGAGAACCCAGTGGAGAGGTTCTCTATCCCTGGCAGGTATCTCTGAACAA 225
Db 1 IleValGlyGlyThrAsnSerSerTrpGlyIleTrpProTrpGlnValSerLeuGlnVal 20
Qy 226 AGG-----CAGAACGATATTGTGGAGAGACATCGTCTCACCACAGTGGGTGATC 276
Db 21 LysLeuThrAlaGlnArgHisLeuCysGlySerLeuIleGlyHisGlnTrpValLeu 40
Qy 277 ACGCGCGCTCACTGCATTGCAACAGAAACATTTGTCTACTTGTGAATGTTACTGCTGGA 336
Db 41 ThrAlaAlaHisCysPheAspGlyLeuProLeuGlnAspValTrpArgIleTyrSerGly 60
Qy 337 GAGTATGACTTAAGCCACAGACAGACCCAGGAGAGAACTCTCAGTATTGAACTGTGATC 396
Db 61 IleLeuAsnLeuSerAspIleThrLysAspThrProPheSerGlnIleLysGluIle 80
Qy 397 ATACATCCACATTTCTCCACCAAGAACCAATGATGATATTCCTTTTGAAGATG 456
Db 81 IleHisGlnAsnTyrLysValSerGluGly---AsnHisAspIleAlaLeuIleLysLeu 99
Qy 457 GCTGGAGCCCTTCCAAATTTGGCCACTTTGTGGGGCCCAATATGTCTTCCAGAGCTCGGGAG 516
```

```

Db 100 GlnAlaProLeuAsnFyrThrGluPheGlnLysProIleCysLeuProSerLysGlyAsp 119
QY 517 CAATTTGAGCGTGTATTTATGACAACTGCAGCGTGGCGCGCTTAACCTGAAGTGGC 576
Db 120 ThrSerThrIleFyrThrAsnCysTrpValThrGlyTrpGlyPheSerLysGluLysGly 139
QY 577 GTCCTCTCAACAGCTTGCAGGAAGTGAATCGCCCTATTGACCTGGGAAGAGTGTGTG 636
Db 140 GluIleGlnAsnIleLeuGlnLysValAsnIleProLeuValThrAsnGluGlnCysGln 159
QY 637 GCAGCTCTGTTAACTAAAGAGGCCATCAGTGGGAAGACCTTCTTTCACACAGCTTTT 696
Db 160 LysArgTyrGlnAspFyrLys-----IleThrGlnArgMetValCysAlaGlyTyr 176
QY 697 CCTGATGGAGGAGACACCATCTCAGGAGATTCAGGAGGTTCACCTCATGTCGCCGAAT 756
Db 177 LysGluGlyCysLysAspLysGlyAspSerGlyCysGlyProLeuValCysLysHis 196
QY 757 AAGAAAGGGCGCTGGACTCTGGCTGTGTGACTTCTCGGGGTTTGGGCTGTGTGTCGAGGC 816
Db 197 ---AsnGlyMetTrpArgLeuValGlyIleThrSerTrpGlyGluCysAlaArg--- 214
QY 817 TGGAGAAACAATGTGAGGAAAAGTGAATCAAGATCCCTCGGGATCTTCACACACATTA 876
Db 215 -----ArgGluGlnProGlyValThrLysValAla 225
QY 877 AAGTGCCTTCTCGGATCCAGCAGACATCCAACTGGTAAC 918
Db 226 GluTyrMetAspTrpIleLeuGluLysThrGlnSerSerAsp 239
RESULT 3
US-09-261-416-2
; Sequence 2, Application US/09261416A
; Patent No. 6291663
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: TADG-12: A No. 6291663el Transmembrane Serine Protease
; TITLE OF INVENTION: Overexpressed in Ovarian Carcinoma
; FILE REFERENCE: D6192
; CURRENT APPLICATION NUMBER: US/09/261,416A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 2
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Amino acid sequence of TADG-12 encoded by nucleotides
; Patent No. 6291663
US-09-261-416-2
Alignment Scores:
Pred. No.: 1,34e-40 Length: 455
Score: 446.50 Matches: 98
Percent Similarity: 49.03% Conservative: 57
Best Local Similarity: 31.01% Mismatches: 112
Query Match: 26.77% Indels: 49
DB: 4 Gaps: 10
US-09-735-713a-1 (1-921) x US-09-261-416-2 (1-455)
QY 1 ATGAGTCTCAAAATGTTATAGCAGGAACAAGCTGATTTTACTACTAGTAATAGTCTTT 60
Db 171 ValSerIleAspHisLeuLeuProAspAspLysValThrAlaLeuHisSerValTyr 190
QY 61 TTTGAACRAGGT---AAATCTGCARTCTTTCGCTCCCAAGCTCCCGAGTGTGGCGAG 117
Db 191 ValArgGluGlyCysAlaSerGlyHisValValThrLeuGlnCysThrAlaCysGlyHis 210
QY 118 AGTCTGTTAAGGTACAGCGCTTGGAAATTATTTTAAACATTTTCAGTCGCATCTTCTTGAGGA 177
```

```

Db 211 ArgArg-----GlyTyrSerSerArgIleValGlyGly 221
QY 178 AGCCAACTGGAGAGGGTTCTATCCCTGGCAGGTATCTCTGAACAACAGGCAAGACAT 237
Db 222 AsnMetSerLeuLeuSerGlnTrpProTrpGlnAlaSerLeuGlnPheGlnGlyTyrHis 241
QY 238 ATTTGTGGAGGAGCATCGTCTCACCAAGTGGGTATCATCAGCGCGCTCAGTCTCATTTGCA 297
Db 242 LeuCysGlyGlySerValIleThrProLeuTrpIleIleThrAlaAlaHisCysValTyr 261
QY 298 ACACAAAACATTTGTCTACTTTGAATGTACTGTCTGGAGAGTATGACTTAAAGCCAGACA 357
Db 262 AspLeuTyrLeuProLysSerTrpThrIleGlnValCysLeuValSerLeuLeuAsp--- 280
QY 358 GACCCAGGAGAGCAAACTCTCAGTATTGAAGTGTCTATCATCATCATCTTCTCCACC 417
Db 281 AsnProAlaProSerHisLeu---ValGluLysIleValTyrHisSerLysTyrLysPro 299
QY 418 AAGAAACCAATGGACTATGATATTGCGCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGC 477
Db 300 LysArg---LeuGlyAsnAspIleAlaLeuMetLysLeuAlaGlyProLeuThrPheAsn 318
QY 478 CACTTTGTGGGGCCCATATGCTTCCAGAGCTCGCGGAGCAATTTTGAAGCTGGTGTATTT 537
Db 319 GluMetIleGlnProValCysLeuProAsnSerGluGluAsnPheProAspGlyLysVal 338
QY 538 TGTACAACTGCAGGCTGGGCGCTTAACCTGAAGTGGGCTCTCTCACAAGTCTTTCAG 597
Db 339 CysTrpThrSerGlyTrpGlyAlaThrGluAspGlyGlyAspAlaSerProValLeuAsn 358
QY 598 GAAGTGAATCTGCTATTGACC-----TGG 624
Db 359 HisAlaAlaValProLeuIleSerAsnLysAspLeuGlnProGlnGlyArgValArgTrp 378
QY 625 GAAGAGTGTGTGGCAGCTCTGTTTAACTAAAGAGGCCCATCATCAGTGGGAGACCTTTCTT 684
Db 379 HisHisLeuProLeuHisAlaLeu----- 386
QY 685 TGCACAGGTTTCTCTGATGAGGGAGA---GACCGATGTGAGGAGATTCAGGAGGTTCA 741
Db 387 ---ArgGlyLeuProAspGlyTrpArgTrpAsnSerCysGlnGlyAspSerGlyGlyPro 405
QY 742 CTCATGTGCGGAATAAGAAAGGGCGCTGAGCTCTGCGTGTGCTGCTCTCCGCGGTTTG 801
Db 406 LeuValCysGlnGluArgArg---LeuTrpLysLeuValGlyAlaThrSerPheGlyIle 424
QY 802 GGCTGTGGTGGCTGGAGAAACAATGTGAGGAAAGTATCAAGGATCCCTCGGATC 861
Db 425 GlyCys-----AlaAspValAsnLysProGlyVal 434
QY 862 TTCACAGACATTAGTAAAGTGTCTTCTGATCCAGCAACACATCCAA 909
Db 435 TyrThrArgValThrSerPheLeuAspTrpIleHisGlnMetGlu 450
RESULT 4
US-08-681-151-3
; Sequence 3, Application US/08681151
; Patent No. 5869637
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Bandman, Olga
; APPLICANT: Braxton, Scott Michael
; APPLICANT: Goli, Surya
; TITLE OF INVENTION: A NOVEL HUMAN KALLIKREIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
```

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/681,151
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0074US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 638 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 205011
; US-08-681-151-3

Alignment Scores:
Pred No.: 1.76e-40 Length: 638
Score: 446.00 Matches: 91
Percent Similarity: 53.12% Conservative: 45
Best Local Similarity: 35.55% Mismatches: 102
Query Match: 26.74% Indels: 18
DB: 2 Gaps: 5

US-09-735-713a-1 (1-921) x US-08-681-151-3 (1-638)
QY 154 ATTTTCAGTCGATCTTGGAGGAGCAAGTGGAGAGGTTCTCTATCCCTGGCAGGTA 213
   ||| :|||||:|||||: ||| :|||||:|||||:
Db 387 IleaSnAlaArgIleValGlyThrAsnSerSerLeuGlyGluTrpProTrpGlnVal 406
   ||| :|||||:|||||: ||| :|||||:|||||:
QY 214 TCTCTGAACAAAGG-----CAGAGCAATTTGTGGAGAAAGCATGCTGCACCA 264
   ||| :|||||:|||||: ||| :|||||:|||||:
Db 407 SerLeuGlnValLysLeuValSerGlnAsnHisMetCysGlyGlySerIleGlyArg 426
   ||| :|||||:|||||: ||| :|||||:|||||:
QY 265 CAGTGGGTGATCAGCGGGCTCACTGTCATTGCAAAACAGAAACATTTGTCTACTTGAAT 324
   ||| :|||||:|||||: ||| :|||||:|||||:
Db 427 GlnTrpIleLeuThrAlaAlaHisCysPheAspGlyIleProTyrProAspValTrpArg 446
   ||| :|||||:|||||: ||| :|||||:|||||:
QY 325 GTTACTCTGGAGATGATGACTTAAGCCAGACAGACCCAGGAGAGCAAACTCTCACTATT 384
   ||| :|||||:|||||: ||| :|||||:|||||:
Db 447 IleTyrGlyGlyIleLeuAsnLeuSerGluIleThrAsnLysThrProPheSerSerIle 466
   ||| :|||||:|||||: ||| :|||||:|||||:
QY 385 GAAACTGTCATCATACATCCACATTTCTCCACCAAGAAACCAATGGACTATGATATGGC 444
   ||| :|||||:|||||: ||| :|||||:|||||:
Db 467 LysGluLeuIleIleHisGlnLysTyrLysMetSerGlyGly---SerTyrAspIleAla 485
   ||| :|||||:|||||: ||| :|||||:|||||:
QY 445 CTTTGTGAAGATGGGTGGAGCCTTCCATTTGGCCACATTTGGGGCCCATATGCTTCCA 504
   ||| :|||||:|||||: ||| :|||||:|||||:
Db 486 LeuIleLysLeuGlnThrProLeuAsnTyrThrGluPheGlnLysProIleCysLeuPro 505
   ||| :|||||:|||||: ||| :|||||:|||||:
QY 505 GAGCTGGCGGAGCAATTTGAGGCTGGTTTATTGTACAACTGACAGCTGGGGCGCGCTTA 564
   ||| :|||||:|||||: ||| :|||||:|||||:
Db 506 SerLysAlaAspThrAsnThrIleTyrThrAsnCysTrpValThrGlyTrpGlyTyrThr 525
   ||| :|||||:|||||: ||| :|||||:|||||:
QY 565 ACTGAAGTGGCGTCTCTCAAGTCTTGCAAGAAAGTGAATCGCTATTTTGAACCTGG 624
   ||| :|||||:|||||: ||| :|||||:|||||:
Db 526 LysGluArgGlyGluThrGlnAsnIleLeuGlnLysAlaThrIleProLeuValProAsn 545
   ||| :|||||:|||||: ||| :|||||:|||||:

;
; US-08-200-900A-2
; Sequence 2, Application US/08200900A
; Patent No. 5665566
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: CLONING OF ENTEROKINASE AND METHOD OF USE
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc. - Legal Affairs
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/200,900A
; FILING DATE: 23-FEB-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meinhert, Maureen C.
; REGISTRATION NUMBER: 31,544
; REFERENCE/DOCKET NUMBER: GI 5201-FWC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 876-1170 X8574
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 798 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-200-900A-2

Alignment Scores:
Pred. No.: 4.69e-40 Length: 798
Score: 442.50 Matches: 96
Percent Similarity: 54.55% Conservative: 54
Best Local Similarity: 34.91% Mismatches: 90
Query Match: 26.53% Indels: 35
DB: 1 Gaps: 9

US-09-735-713a-1 (1-921) x US-08-200-900A-2 (1-798)
QY 106 AGTTGTGGCAGAGCTGTGGTT-----AAGTACAGCCTTGGAAATTATTTAACATTTC 159
   |||||||: ||| :|||||: |||
```

```
Db 550 SerCysGlyLysLysLeuValThrGlnGluValSerPro-----562
Qy 160 AGTCGCATCTTGGAGGAGCCCAAGTGGAGAGGTTCCCTATCCTCCGCGAGGTATCTCTG 219
Db 563 ---LysIleValGlySerAspSerArgGluGlyAlaTrpProTrpValValAlaLeu 581
Qy 220 AAACAAAGGCGAAGCATATTTGGAGGAGCATCTCTCACCACAGTGGGTGATCAGC 279
Db 582 TyrPheAspGlnGlnValCysGlyAlaSerLeuValSerArgAspTrpLeuValSer 601
Qy 280 GCGCTCCTCCTCAATTTGCAACAGAAACATT---GTGCTCTACTTTGAATGTTACTGCTGGA 336
Db 602 AlaAlaHisCysValTyrglyArgAsnMetGluProSerLysTrpLysAlaValLeuGly 621
Qy 337 GAGTATGACTTAAGCCAG---ACAGACCCAGGAGAGCAAACTCTCATTATGAACATGTC 393
Db 622 LeuHisMetAlaSerAsnLeuThrSerProGlnIleGluThrArgLeuIleAspGlnIle 641
Qy 394 ATCATACATCCACATTTCTCCACCAAGAAACCAATGGACTATGATATTGTCCTTTGAAG 453
Db 642 ValIleAsnProHisTyrAsn---LysArgArgLysAsnAsnAspIleAlaMetMetHis 660
Qy 454 ATGGCTGGAGCCTTCCAAATTTGGCCACTTTGTGGGCCCATPATCTCTTCCAGAGCTCGG 513
Db 661 LeuGluMetLysValAsnTyrThrAspTyrIleGlnProIleCysLeuProGluGluAsn 680
Qy 514 GAGCAATTTGAGGTGGTTTATTGTACAACTGCAGCTCGCGGCGCTTAACTGAAGGT 573
Db 681 GlnValPheProGlyArgIleCysSerIleAlaGlyTrpGlyAlaLeuIleTyrGln 700
Qy 574 GCGCTCTCTCACAAGTCTTCCAGGAGTGAATCTGCTATTTTGCCTGAGAGAGTGT 633
Db 701 GlySerThrAlaAspValLeuGlnGluAlaAspValProLeuLeuSerAsnGluLysCys 720
Qy 634 GTGCAGCTCTG-----TTAACACTAAAGAGGCCCATCAGTGGGAAGACCTTT 681
Db 721 GlnGlnMetProGluTyrAsnIleThr-----GluAsnMet 733
Qy 682 CTTTGCAGAGTTTCTCGATGGAGGAGAGACGCATGTCAGGAGATTCAGGAGGTTC 741
Db 734 ValCysAlaGlyTyrGluAlaGlyValAspSerCysGlnGlyAspSerGlyGlyPro 753
Qy 742 CTCTATGTCGCGGAATAAGAAAGGCGCTCTGCTGGTGTGTGCTCTCTGCGGTTG 801
Db 754 LeuMetCysGln---GluAsnAsnArgTrpLeuLeuAlaGlyValThrSerPheGlyTyr 772
Qy 802 GCGTGTGCTCAGGCTGGAGAAACATGTGAGGAAAAGTGTATCAAGGATCCCTGGGATC 861
Db 773 GlnCysAlaLeuProAsnArg-----ProGlyVal 782
Qy 862 TTCACAGACATTAGTAAAGTCTTCTCTGGATCCACGACACATC 906
Db 783 TyrAlaArgValProArgPheThrGluTrpIleGlnSerPheLeu 797

RESULT 6
PCT-US94-00616-2
; Sequence 2, Application PC/TUS9400616
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: CLONING OF ENTEROKINASE AND METHOD OF USE
; NUMBER OF SEQUENCES: 33
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/00616
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 798 amino acids
```

```
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-00616-2

Alignment Scores:
Pred. No.: 4, 69e-40 Length: 798
Score: 442.50 Matches: 96
Percent Similarity: 54.55% Conservative: 54
Best Local Similarity: 34.91% Mismatches: 90
Query Match: 26.53% Indels: 35
DB: 5 Gaps: 9

US-09-735-713a-1 (1-921) x PCT-US94-00616-2 (1-798)
Qy 106 AGTTGTGGGCGAGCTCTGGTT-----AAGCTACAGCCTTGGAAATTATTTAAACATTTTC 159
Db 550 SerCysGlyLysLysLeuValThrGlnGluValSerPro-----562
Qy 160 AGTCGCATCTTGGAGGAGCCCAAGTGGAGAGGTTCCCTATCCTCCGCGAGGTATCTCTG 219
Db 563 ---LysIleValGlySerAspSerArgGluGlyAlaTrpProTrpValValAlaLeu 581
Qy 220 AAACAAAGGCGAAGCATATTTGGAGGAGCATCTCTCACCACAGTGGGTGATCAGC 279
Db 582 TyrPheAspGlnGlnValCysGlyAlaSerLeuValSerArgAspTrpLeuValSer 601
Qy 280 GCGCTCCTCCTCAATTTGCAACAGAAACATT---GTGCTCTACTTTGAATGTTACTGCTGGA 336
Db 602 AlaAlaHisCysValTyrglyArgAsnMetGluProSerLysTrpLysAlaValLeuGly 621
Qy 337 GAGTATGACTTAAGCCAG---ACAGACCCAGGAGAGCAAACTCTCATTATGAACATGTC 393
Db 622 LeuHisMetAlaSerAsnLeuThrSerProGlnIleGluThrArgLeuIleAspGlnIle 641
Qy 394 ATCATACATCCACATTTCTCCACCAAGAAACCAATGGACTATGATATTGTCCTTTGAAG 453
Db 642 ValIleAsnProHisTyrAsn---LysArgArgLysAsnAsnAspIleAlaMetMetHis 660
Qy 454 ATGGCTGGAGCCTTCCAAATTTGGCCACTTTGTGGGCCCATPATCTCTTCCAGAGCTCGG 513
Db 661 LeuGluMetLysValAsnTyrThrAspTyrIleGlnProIleCysLeuProGluGluAsn 680
Qy 514 GAGCAATTTGAGGTGGTTTATTGTACAACTGCAGCTCGCGGCGCTTAACTGAAGGT 573
Db 681 GlnValPheProGlyArgIleCysSerIleAlaGlyTrpGlyAlaLeuIleTyrGln 700
Qy 574 GCGCTCTCTCACAAGTCTTCCAGGAGTGAATCTGCTATTTTGCCTGAGAGAGTGT 633
Db 701 GlySerThrAlaAspValLeuGlnGluAlaAspValProLeuLeuSerAsnGluLysCys 720
Qy 634 GTGCAGCTCTG-----TTAACACTAAAGAGGCCCATCAGTGGGAAGACCTTT 681
Db 721 GlnGlnMetProGluTyrAsnIleThr-----GluAsnMet 733
Qy 682 CTTTGCAGAGTTTCTCGATGGAGGAGAGACGCATGTCAGGAGATTCAGGAGGTTC 741
Db 734 ValCysAlaGlyTyrGluAlaGlyValAspSerCysGlnGlyAspSerGlyGlyPro 753
Qy 742 CTCTATGTCGCGGAATAAGAAAGGCGCTCTGCTGGTGTGTGCTCTCTGCGGTTG 801
Db 754 LeuMetCysGln---GluAsnAsnArgTrpLeuLeuAlaGlyValThrSerPheGlyTyr 772
Qy 802 GCGTGTGCTCAGGCTGGAGAAACATGTGAGGAAAAGTGTATCAAGGATCCCTGGGATC 861
Db 773 GlnCysAlaLeuProAsnArg-----ProGlyVal 782
Qy 862 TTCACAGACATTAGTAAAGTCTTCTCTGGATCCACGACACATC 906
Db 783 TyrAlaArgValProArgPheThrGluTrpIleGlnSerPheLeu 797

RESULT 7
```

	QY	406	CATTTC-----TCCACCAAGAAACAATGGACTATGCTATGTATTTGCCCTTTTG	450
	Dd	241	GlyThrLeuProPheArgAspProThrIleaspGluasnSerasnAspIleAlaLeuVal	260
	QY	451	AAGATGGCTGGAGCCTTCCAATTGGCGCACCTTTGTGGGGCCCATAATGTCTTCCAGAGCTG	510
	Dd	261	HisLeuSerSerLeuProLeuThrGluTyrlleGlnProValCysLeuProAlaLa	280
	QY	511	CGGAGCAATTTGAGGCTGGTTTTATTGTACAACTCCAGGCCTGGGGCCGCTTAACCTGA	570
	Dd	281	GlyGlnAlaLeuValAspGlyLysValCysThrValThrGlyTrpGlyAsnThrGlnPhe	300
	QY	571	GGTGGCGTCCTCTCACAACTTCAGAGAAGTGAATCTGCCTATTMTTAGCTGGGAAGAG	630
	Dd	301	TyrGlyGlnGlnAlaMetValLeuGlnGluAlaArgValProIlelleSerasnGluVal	320
	QY	631	TGTGTGGCAGCTCTGTTAACACTAAGAGGCC--	672
	Dd	321	Cys-----AsnSerProAspPheTyrglyAsnGlnIleLysPro	333
	QY	673	AAGACCTTCTTCACAGAGTTTTCNGATGGAGGGAGAGACCAGCATCTCAGGAGATTCA	732
	Dd	334	LysMetPhe---CysAlaGlyTyrrProgluglyIleaspaLcysGlnGlyAspSer	352
	QY	733	GGAGGTTCACTCATGTGCCCGAATAAG- - - - - AAAGGGCCTGGACTCTGGCTGCT	783
	Dd	353	GlyGlyProPheValCysGluAspSerIleSerGlyThrSerArgTrpargLeuCysGly	372
	QY	784	GTGACTTCCTGGGGTTTGGCCTGTGTCGAGGCTGGAGAACAACTGTAGGAAAAAGTGAT	843
	Dd	373	IleValSerTrpGlyThrGlyCysAlaLeu- - - - - AlaargLys-----	385
	QY	844	CAAGGATCCCCCTGGGATCTTCACAGACATAGTAAAAGTCTTTCCTGGATCCACGAACAC	903
	Dd	386	- - - - - ProGlyValTyrrLysValThrAspPheargGluTrpillePhelysAla	402
	QY	904	ATCCAAACT 912	
	Dd	403	IleLysThr 405	
	RESULT	8		
	US-09-386-642-11			
	; Sequence 11, Application US/093866642			
	; Patent No. 6420157			
	; GENERAL INFORMATION:			
	; APPLICANT: Darow, Andrew			
	; APPLICANT: Qi, Jensen			
	; APPLICANT: Andrade-Gordon, Patricia			
	; TITLE OF INVENTION: Zymogen Activation System			
	; FILE REFERENCE: ORT-1028			
	; CURRENT APPLICATION NUMBER: US/09/386,642			
	; CURRENT FILING DATE: 1999-08-31			
	; NUMBER OF SEQ ID NOS: 60			
	; SOFTWARE: PatentIn Ver. 2.0			
	; SEQ ID NO 11			
	; LENGTH: 328			
	; TYPE: PRT			
	; ORGANISM: Artificial Sequence			
	; FEATURE:			
	; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene			
	; OTHER INFORMATION: with homo sapien serine protease catalytic domain			
	US-09-386-642-11			
	Alignment Scores:			
	Pred. No.: 6.88e-40 Length: 328			
	Score: 439.50 Matches: 99			
	Percent Similarity: 48.71% Conservative: 52			
	Best Local Similarity: 31.94% Mismatches: 120			
	Query Match: 26.35% Indels: 39			
	DB: 4 Gaps: 8			
	US-09-735-713A-1 (1-921) x US-09-386-642-11 (1-328)			

```
QY 7 CTCAAAATGCTTATAAGCAGGAAGCTGATTTTACTACTAGGAATAGTCTTT----- 60
Db 12 LeuLeuLeuLeuValSerAsnLeuLeuLeuLeuLeuLeuValValSerAspTyr 31
QY 61 -----TTTGAACRAGGTAAATCGARCTCTTTTCGCTCCCAAGCTCCAGTGTGGG 114
Db 32 LysAspAspAspValAspAlaAlaLeuAlaAlaPro----- 45
QY 115 CAGAGCTGTTAAGCTACAGCCTTGAATATTATTTAAACATTTTCAGTCGCATCTTGA 174
Db 46 -----PheAspAspAspLysIleValGly 54
QY 175 GGAAGCAAGTGGAGAAGGTTCTCTATCCCTGGCAGCTATCTCTGAACAAAGGCAGAAG 234
Db 55 GlyTyrAlaLeuLeuAlaGlyGlnTyrProTyrGlnValSerIleThrTyrGluGlyVal 74
QY 235 CATATTGTTGGAGGAGCATCGCTCCACACAGTGGGTGATCAGCGGCTCACTGCATT 294
Db 75 HisValCysGlySerLeuValSerGluGlnTyrValLeuSerAlaAlaHisCysPhe 94
QY 295 GCAACAGAAACATCTGCTACTTGAATGTTACTGCTGGAGATGACTTTAAGCCAG 354
Db 95 ProSerGluHisLysGluAlaTyrGluValLysLeuGlyAlaHisGlnLeuAspSer 114
QY 355 ACAGACCCAGGAGCAAACTCTCACTATTGAACTGTATCATATACATCCACATTTCTCC 414
Db 115 TyrSerGluAspAlaLysValSerThrLeuLysAspIleProHisProSerTyr--- 133
QY 415 ACCAAGAACCAATGACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 474
Db 134 LeuGlnGluGlySerGlnGlyAspIleAlaLeuLeuGlnLeuSerArgProIleThrPhe 153
QY 475 GGCCACTTGTGGGCCCCATATCTCTCCAGACCTCGGGAGCAATTTGAGGCTGCTTT 534
Db 154 SerArgTyrIleArgProLysCysLeuProAlaAlaAsnAlaSerPheProAsnGlyLeu 173
QY 535 ATTTGTACAACTCAGGCTGGGCGCCCTTAACCTGAAGGTGGCGTCTCTC----- 588
Db 174 HisCysThrValThrGlyTyrGlyHisValAlaProSerValSerLeuLeuThrProLys 193
QY 589 GTCTTCAGGAAGTGAATCGCTATTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTGA 648
Db 194 ProLeuGlnGlnLeuValProLeuIleSerArgGluThrCysAsnCysLeuTyrAsn 213
QY 649 ACATAAAGAGGCC-----ATCAGTGGGAAGACCTTTCTTTGTCACAGGTTT 696
Db 214 IleAspAlaLysProGluGluProHisPheValGlnGluAspMetValCysAlaGlyTyr 233
QY 697 CTGTATGGAGGAGAGACCATGTCAGGAGATTCAGGAGGTTTCACTCATGTGCGCGAAT 756
Db 234 ValGluGlyLysAspAlaCysGlnGlyAspSerGlyGlyProLeuSerCys---Pro 252
QY 757 AAGAAAGGCGCTGGACTCGGCTGGTGTGACTTCTCGGGTGTGGCTGTGGCTGAGGC 816
Db 253 ValGluGlyLeuTyrTyrLeuThrGlyIleValSerTyrProGlyAspAlaCysGlyAla--- 271
QY 817 TGCAGAAACAAATGTGAGGAAAGTCAAGGATCCCTCGGATCTTCACAGACATTAGT 876
Db 272 ---ArgAsnArg-----ProGlyValTyrThrLeuAlaSer 282
QY 877 AAAGTCTTTCTCGGATCCACGACACATC 906
Db 283 SerTyrAlaSerTrpIleGlnSerLysVal 292
```

## RESULT 9

US-09-027-337-3  
; Sequence 3, Application US/09027337B  
; Patent No. 5972616  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Tanimoto, Hiroto  
; TITLE OF INVENTION: TAGD-15: An Extracellular Serine Protease Overexpressed in  
; TITLE OF INVENTION: Breast and Ovarian Carcinomas

```
; FILE REFERENCE: D6064  
; CURRENT APPLICATION NUMBER: US/09/027,337B  
; CURRENT FILING DATE: 1998-02-20  
; NUMBER OF SEQ ID NOS: 13  
; SEQ ID NO 3  
; LENGTH: 256  
; TYPE: PRT  
; ORGANISM: Unknown  
; FEATURE:  
; OTHER INFORMATION: Serine protease catalytic domain of hepsin (Heps)  
; OTHER INFORMATION: homologous to similar domain in TAGD-15  
US-09-027-337-3
```

Alignment Scores: 1.03e-39 Length: 256  
Pred. No.: 437.50 Matches: 93  
Score: 51.35% Conservative: 40  
Percent Similarity: 35.91% Mismatches: 103  
Best Local Similarity: 26.23% Indels: 23  
Query Match: 2 Gaps: 7  
DB:

US-09-735-713a-1 (1-921) x US-09-027-337-3 (1-256)

```
QY 163 CGCATTTCTGGAGGAAGCCAAAGTGGAGAAGGTTCTTATCCCTGGCAGGTATCTCGAA 222
Db 1 ArgIleValGlyGlyArgAspThrSerLeuGlyArgTyrProTyrGlnValSerLeuArg 20
QY 223 CAAAGCCAGAGCATATTTGTGGAGGAGCATGCTCCACACAGTGGGTGATCAGCGG 282
Db 21 TyrAspGlyAlaHisLeuCysGlyGlySerLeuLeuSerGlyAspThrValLeuThrAla 40
QY 283 GCTCACTGATGTCACAAACAGAAAC---ATTGTGCTCTACTTTGAATGTTACTGCTGGAGAG 339
Db 41 AlaHisCysPheProGluArgAsnArgValLeuSerArgTyrPheValPheAlaGlyAla 60
QY 340 TATGACTTAAGCCAGACAGACCCAGGAGAGCAAACTCTCACTATTGAAACTGTCATCAT 399
Db 61 -----ValAlaGlnAlaSerProHisGlyLeuGlnGlyValGlnAlaValTyr 78
QY 400 CATCCACATTTTC-----TCCACCAAGAAACCAATGACTATGATATTGCC 444
Db 79 HisGlyGlyTyrLeuProPheArgAspProAsnSerGluGluAsnSerAsnAspIleAla 98
QY 445 CTTTGAAGATGGCTGGAGCCTTCCAAATTTGGCCACTTTGTGGGCGCCATATCTCTCCA 504
Db 99 LeuValHisLeuSerSerProLeuProLeuThrGluTyrIleGlnProValCysLeuPro 118
QY 505 GAGCTCGGGGCAATTTGAGGCTGTTTATTTTGTACAACTGCAGGCTGGGCGCGCTTA 564
Db 119 AlaAlaGlyGlnAlaLeuValAspGlyLysIleCysThrValThrGlyTyrGlyAsnThr 138
QY 565 ACTGAAGTGGCGTCTCTCACAAAGTCTTCCAGGAAGTGAATGCTGCTATTGACCTGG 624
Db 139 GlnTyrTyrGlyGlnGlnAlaGlyValLeuGlnGluAlaArgValProIleLeuSerAsn 158
QY 625 GAAGAGTGTGTGGCAGCTCTGTTAAACACTAAAGAGGCCCATCAGTGGGAAGACCTTCTT 684
Db 159 AspValCysAsnGlyAlaAspPhe---TyrGlyAsnGlnIleLysProLysMetPhe--- 176
QY 685 TGCACAGGTTTCTCTGATGGAGGAGAGACGTCATGTCAGGGAGATTCAGAGGTTCACTC 744
Db 177 CysAlaGlyTyrProGluGlyGlyIleAspAlaCysGlnGlyAspSerGlyGlyProPhe 196
QY 745 ATGTGCGCG-----AATAAGAAAGGGCGCTGGACTGCTGGTGGTGTGACTTCTGG 795
Db 197 ValCysGluAspSerIleSerArgThrProArgTyrProArgLeuCysGlyIleValSerTrp 216
QY 796 GGTTTGGGCTGTGCTCGAGGCTGGAGAAACAATGTGAGGAAAGTGTATCAAGGATCCCT 855
Db 217 GlyThrCysAlaLeuAlaGlnLys-----Pro 226
QY 856 GGGATCTTCACAGACATTAGTAAGATGCTTCTCTGGATCCACGAACACATCCAACCT 912
Db 856 GGGATCTTCACAGACATTAGTAAGATGCTTCTCTGGATCCACGAACACATCCAACCT 912
```

27

Db 227 GlyValTyrThrLysValSerAspPheArgGluTrpIlePheGlnAlaIleLysThr 245  
RESULT 10  
US-09-644-600-3  
; Sequence 3, Application US/09644600  
; Patent No. 6451500  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, Timothy J.  
; APPLICANT: Tanimoto, Hiroto  
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease  
; TITLE OF INVENTION: Overexpressed in Carcinomas  
; FILE REFERENCE: D6064CIE/D  
; CURRENT APPLICATION NUMBER: US/09/644,600  
; CURRENT FILING DATE: 2000-08-23  
; PRIOR APPLICATION NUMBER: 09/421,213  
; PRIOR FILING DATE: 1999-10-20  
; PRIOR APPLICATION NUMBER: 09/027,337  
; PRIOR FILING DATE: 1998-02-20  
; NUMBER OF SEQ ID NOS: 98  
; SEQ ID NO 3  
; LENGTH: 256  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: Hepsin  
US-09-644-600-3  
Alignment Scores:  
Pred. No.: 1.03e-39 Length: 256  
Score: 437.50 Matches: 93  
Percent Similarity: 51.35% Conservative: 40  
Best Local Similarity: 35.91% Mismatches: 103  
Query Match: 26.23% Indels: 23  
DB: 4 Gaps: 7  
US-09-735-713A-1 (1-921) x US-09-644-600-3 (1-256)  
QY 163 CGCATTCCTGGAGGAGCAAGTGGAGAGGTTCTCTCCCTGGCAGGTATCTCTGAAA 222  
Db 1 ArgIleValGlyGlyArgAspThrSerLeuGlyArgTrpProTrpGlnValSerLeuArg 20  
QY 223 CAAGGCGAAGCATATTGTGGAGGAGCATCTCTACACAGTGGGTGATCAGCGCG 282  
Db 21 TyrAspGlyAlaHisLeuCysGlyGlySerLeuLeuSerGlyAspTrpValLeuThrAla 40  
QY 283 GCTCACTGCATTCGAACACAGAAC--ATTGTGTCTACTTTGATGTACTGCTGGAGAG 339  
Db 41 AlaHisCysPheProGluArgAsnArgValLeuSerArgTrpArgValPheAlaGlyAla 60  
QY 340 TATGACTTAAGCCAGACAGACCCAGGAGAGCAAACTCTCACTATTGAAACTGTATCATATA 399  
Db 61 -----ValAlaGlnAlaSerProHisGlyLeuGlnLeuGlyValGlnAlaValValTyr 78  
QY 400 CATCCACATTC-----TCCACCAAGAAACCAATGGACTATGATATTGCC 444  
Db 79 HisGlyGlyTyrLeuProPheArgAspProAsnSerGluGluAsnSerAsnAspIleAla 98  
QY 445 CTTTGTGAAGATGGCTGAGCGCTTCCATTTGGCCACATTTGTGGGGCCCATATGCTTCCA 504  
Db 99 LeuValHisLeuSerSerProLeuProLeuThrGluTyrIleGlnProValCysLeuPro 118  
QY 505 GAGCTGGGGAGCAATTTGAGGCTGGTTTATTGTTTACAACTGCAGCGCTGGGGCCGCTTA 564  
Db 119 AlaAlaGlyGlnAlaLeuValAspGlyLysIleCysThrValThrGlyTrpGlyAsnThr 138  
QY 565 ACTGAAGTGGCTCTCTCACAAGTCTTCAGGAAGTGAATCGCCTATTATTCACCTGG 624  
Db 139 GlnTyrTyrGlyGlnGlnAlaGlyValLeuGlnAlaArgValProIleIleSerAsn 158  
QY 625 GAACAGTGTGTGGCAGCTCTGTTAACTAAAGAGGCCCATCAGTGGGAAGACCTTTCTT 684  
Db 159 AspValCysAsnGlnAlaAspPhe---TyrGlyAsnGlnIleLysProLysMetPhe--- 176

QY 685 TGCACAGGTTTCTCTGATGGAGGAGAGACGCATGTCCAGGAGATTCAGGAGTTCACCT 744  
Db 177 CysAlaGlyTyrProGluGlyGlyIleAspAlaCysGlnGlyAspSerGlyGlyProPhe 196  
QY 745 ATGTGCGCG-----AATAGAAAGGGCTCTGGACTCTGGCTGTGTGACTTCTCTGG 795  
Db 197 ValCysGluAspSerIleSerArgTrpProArgTrpArgLeuGlyGlyIleValSerTrp 216  
QY 796 GGTTTGGCTGTGCTCGAGGCTGGAGAAACAATGTGAGGAAAGTGTATCAAGGATCCCT 855  
Db 217 GlyThrGlyCysAlaLeuAlaGlnLys-----pro 226  
QY 856 GGATCTCTCACAGACATTAGTAAGTCTTCTCTGGATCCACGAAACATCCAAACT 912  
Db 227 GlyValTyrThrLysValSerAspPheArgGluTrpIlePheGlnAlaIleLysThr 245  
RESULT 11  
US-09-386-642-12  
; Sequence 12, Application US/09386642  
; Patent No. 6420157  
; GENERAL INFORMATION:  
; APPLICANT: Darrow, Andrew  
; APPLICANT: Qi, Jensen  
; APPLICANT: Andrade-Gordon, Patricia  
; TITLE OF INVENTION: Zymogen Activation System  
; FILE REFERENCE: ORT-1028  
; CURRENT APPLICATION NUMBER: US/09/386,642  
; CURRENT FILING DATE: 1999-08-31  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 319  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene  
; OTHER INFORMATION: with homo sapien serine protease catalytic domain  
US-09-386-642-12  
Alignment Scores:  
Pred. No.: 1.65e-39 Length: 319  
Score: 436.00 Matches: 90  
Percent Similarity: 51.74% Conservative: 44  
Best Local Similarity: 34.75% Mismatches: 107  
Query Match: 26.14% Indels: 18  
DB: 4 Gaps: 6

US-09-735-713A-1 (1-921) x US-09-386-642-12 (1-319)  
QY 148 TTTAAACATTTTCAGTCGCATTTCTGGAGGAGCAAGTGGAGAGGTTCTCTATCCCTGG 207  
Db 37 PheAspAspAspAspLysIleValGlyGlyTyrAlaLeuGluAlaGlyGlnTrpProTrp 56  
QY 208 CAGTATCTCTGAACAAAGGAGACCATATTGTGGAGGAGCATCTGCTACACAG 267  
Db 57 GlnValSerIleThrTyrGluGlyValHisValCysGlyGlySerLeuValSerGluGln 76  
QY 268 TGGTGTATCAGCGGCTCAGTCATTTGCAACAGAAACATTTGTCTACTTTGAATGTT 327  
Db 77 TrpValLeuSerAlaAlaHisCysPheProSerGluHisLysGluAlaTyrGluVal 96  
QY 328 ACTGCTGGAGATGATGACTTAAGCCAGACAGACCCAGGAGCAAACTCTCACTATTGAA 387  
Db 97 LysLeuGlyAlaHisGlnLeuAspSerTyrSerGluAspAlaLysValSerThrLeuLys 116  
QY 388 ACTGTCTATCATATCCACATTTCTCCACCAAGAAACCAATGGACTATGATATTGCCCTT 447  
Db 117 AspileIleProHisProSerTyr---LeuGlnGluGlySerGlnGlyAspIleAlaLeu 135  
QY 448 TTGAAGATGCTGGAGCCTTCCAATTTGGCCACTTTTGGGGCCCATATCTGTCCAGAG 507  
Db 136 LeuGlnLeuSerArgProIleThrPheSerArgTrpIleArgProIleCysLeuProAla 155



Patent No. 6177258  
GENERAL INFORMATION:  
APPLICANT: Wu Hunter, Shirley  
APPLICANT: Stiegler, Gary  
APPLICANT: Gaines, Patrick J.  
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID  
TITLE OF INVENTION: MOLECULES AND USES THEREOF  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross P.C.  
STREET: 1700 Lincoln Street, Suite 3500  
CITY: Denver  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80203  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/004,731  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/749,699  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Connell, Gary J.  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2618-25-C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 30:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 400 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-004-731-30

Alignment Scores:  
Pred. No.: 4,4e-39 Length: 400  
Score: 432.50 Matches: 101  
Percent Similarity: 53.85% Conservative: 46  
Best Local Similarity: 37.00% Mismatches: 91  
Query Match: 25.93% Indels: 35  
DB: 4 Gaps: 11

US-09-735-713A-1 (1-921) x US-09-004-731-30 (1-400)

QY 106 AGTTGTGGCAGAGCTGGTTAAGTACAGCCTTGGAAATTTTACATTTTCAGTCGC 165  
Db 148 ThrCysGlyGlu-----LeuYrThrArgSerAsnArg 158  
QY 166 ATTCTTGGAGGAGCCAAAGTGGACAGAGGTTTCCTATCCCTGGCAGGATCTCTCAACAA 225  
Db 159 lleValGlyGlyHisSerThrGlyPheGlySerHisProTrpGlnAlaAlaLeuIleLys 178  
QY 226 -----AGGCAGAGCATATTGTGGAGGAGCAGTCGTCTCACCACAGTGGGTG 273  
Db 179 SerGlyPheLeuSerLysLysLeuSerCysGlyGlyAlaLeuValSerAspArgTrpVal 198  
QY 274 ATCAGCGGGCTCACTCATTCGAAACAGAACATTTGTCTACTTTGAATGTTACTGCT 333  
Db 199 lleThrAlaAlaHisCysValAlaThr---ThrProAsnSerAsnLeuLysValArgLeu 217  
QY 334 GGAGATGATGACTTAAGCCAGCACAGACCCA-----GGAGAGCAAACTCTCAGTATTGAA 387  
Db 218 GlyGluTrpAspValArgAspHisAspGluArgLeuAsnHisGluGluTyAlaIleGlu 237  
QY 388 ACTGTCATCATCATCCACATTTCTCCACCAAGAAACCAATGGACTAT-----GATATT 441

Db 238 ArgLysGluValHisProSerTyrSer-----ProThrAspPheArgAsnAspVal 254  
QY 442 GCCCTTTTGAAGATGCTGGAGCCTTCCAATTTGGCCACTTTTGGGGCCCATATGTCCTT 501  
Db 255 AlaLeuValLysLeuAspArgThrValIlePheLysGlnHisIleLeuProValCysLeu 274  
QY 502 CCAGAGCTGCGGGAGCAATTGGAGCTGTTTATTATTGTACAACCTGAGGCTGGGGCCGC 561  
Db 275 ProHis---LysGlnMetLysLeuAlaGlyLysMetAlaThrValAlaGlyTrpGlyArg 293  
QY 562 TTAACCTGAAGGT---GGCGTCTCTCACAAAGTCTTCGAGGAAGTGAATCTCCCTATTTTG 618  
Db 294 ThrArgHisGlyGlnSerThrValProAlaValLeuGlnGluValAspValGluValIle 313  
QY 619 ACCTGGGAAGAGTGTGGCAGCTCTGTTAACTACTA---AAGAGGCCCATCAGTGGGAAG 675  
Db 314 ProAsnGluArgCysGlnArgTrpPheArgAlaAlaGlyArgGluThrIleHisAsp 333  
QY 676 ACCTTTCTTTGCACAGCTTTTCTGATGGAGGAGGAGACGACATGTCAGGAGATTCAGGA 735  
Db 334 ValPheLeuCysAlaGlyTyLysGluGlyArgAspSerCysGlnGlyAspSerGly 353  
QY 736 GGTTCATCTATGTCCGGAATAAGAAAGGGCCCTGGACTCTGCTGGTGTGTGACTCTCTCG 795  
Db 354 GlyProLeuIleMetGln---IleGluGlyArgArgThrLeuValGlyLeuValSerTrp 372  
QY 796 GGTTTGGCTGTGGTCAGGCTGGAGAAACAATGTCAGGAAAGTGTATCAAGGATCCCT 855  
Db 373 GlyIleGlyCysGlyArg-----GluHisLeuPro 382  
QY 856 GGGATCTTCACAGACATTAGTAAAGTCTTTCCTGGATC 894  
Db 383 GlyValTyThrAsnIleGlnLysPheIleProTrpIle 395

## RESULT 14

US-09-004-731-33  
Sequence 33, Application US/09004731  
Patent No. 6177258  
GENERAL INFORMATION:  
APPLICANT: Wu Hunter, Shirley  
APPLICANT: Stiegler, Gary  
APPLICANT: Gaines, Patrick J.  
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID  
TITLE OF INVENTION: MOLECULES AND USES THEREOF  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross P.C.  
STREET: 1700 Lincoln Street, Suite 3500  
CITY: Denver  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80203  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/004,731  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/749,699  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Connell, Gary J.  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2618-25-C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 33:

SEQUENCE CHARACTERISTICS:  
LENGTH: 400 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-004-731-33

Alignment Scores:  
Pred. No.: 4,4e-39 Length: 400  
Score: 432.50 Matches: 101  
Percent Similarity: 53.85% Conservative: 46  
Best Local Similarity: 37.00% Mismatches: 91  
Query Match: 25.93% Indels: 35  
DB: 11 Gaps: 11

US-09-735-713A-1 (1-921) x US-09-004-731-33 (1-400)

```
QY 106 AGTTGGGCGAGAGTCTGTTAAGGTACAGCCTTGAATTTTAAACATTTTCACTGCGC 165
Db 148 ThrCysGlyGlu-----LeuTyrThrArgSerAsnArg 158
QY 166 ATTCTGGAGGACCAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAACAA 225
Db 159 IleValGlyGlyHisSerThrGlyPheGlySerHisProtrpGlnAlaLeuLys 178
QY 226 -----AGGCAGGAAGCATATTGTGGAGGAGCATCTCTCACCACAGTGGTG 273
Db 179 SerGlyPheLeuSerLysLysLeuSerCysGlyGlyAlaLeuValSerAspArgTrpVal 198
QY 274 ATCAGCGCGCTCAGTCATGTCGAATGCAACAGAACATTGTGCTACTTTGAATGTACTGCT 333
Db 199 IleThrAlaAlaHisCysValAlaThr---ThrProAsnSerAsnLeuLysValArgLeu 217
QY 334 GGAGATATGACTATAGCCAGACAGACCCCA-----GGAGAGCAAACTCTCATTATTGAA 387
Db 218 GlyGluTrpAspValArgAspHisAspGluArgLeuAsnHisGluLufyAlaIleGlu 237
QY 388 ACTGTCATCATCATACATCCACATTTCTCCACCAAGAAACATGAGCATAT-----GATATT 441
Db 238 ArgLysGluValHisProSerTyrSer-----ProThrAspPheArgAsnAspVal 254
QY 442 GCCCTTTTGAAGTGGCTGGAGGCTTCCAAATTTGGCCACTTTGTGGGCCCATATGCTTT 501
Db 255 AlaLeuValLysLeuAspArgThrValIlePheLysGlnHisIleLeuProValCysLeu 274
QY 502 CCAGAGCTGGCGGAGCAATTTGAGGTGGTATTATTTGTACAACTCGAGGCTGGGGCCGC 561
Db 275 ProHis---LysGlnMetLysLeuAlaGlyLysMetAlaThrValAlaGlyTrpGlyArg 293
QY 562 TTAAGTGAAGGT---GGCGCTCTCACAGTCTTGCAGGAGTGTGCGCTATTTTG 618
Db 294 ThrArgHisGlyGlnSerThrValProAlaValLeuGlnGluValAspValGluValIle 313
QY 619 ACCTGGGAAGAGTGTGGCAGCTCTGTTAACACTA---AAGAGGCCCATCAGTGGGAAG 675
Db 314 ProAsnGluArgCysGlnArgTTPheArgAlaAlaGlyArgGluThrIleHisAsp 333
QY 676 ACCTTTCTTTCACAGGTTTCTGATGGAGGAGACGCCATGTCAGGAGATTGAGGA 735
Db 334 ValPheLeuCysAlaGlyTyrLysGluGlyArgAspSerCysGlnGlyAspSerGly 353
QY 736 GGTTCCTACTATGTCGCGGAATAAGAAGGCGCTGAGCTGCTGCTGCTGCTGCTGCTG 795
Db 354 GlyProLeuIleMetGln---IleGluGlyArgArgThrLeuValGlyLeuValSerTrp 372
QY 796 GGTTCGCTGCTGCTGAGGCTGGAGAAACAATGTGAGGAAAAAGTGATCAAGGATCCCT 855
Db 373 GlyIleGlyCysGlyArg-----GluHisLeuPro 382
QY 856 GGATCTTCACACACATAGTAAGTCTTCTCTGGATC 894
Db 383 GlyValTyrThrAsnIleGlnLysPheIleProTrpIle 395
```

RESULT 15

US-08-749-699-30  
Sequence 30, Application US/08749699  
Patent No. 6210920  
GENERAL INFORMATION

APPLICANT: Wu Hunter, Shirley  
APPLICANT: Stiegler, Gary  
APPLICANT: Gaines, Patrick J.  
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID  
MOLECULES AND USES THEREOF  
NUMBER OF SEQUENCES: 103

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross P.C.

STREET: 1700 Lincoln Street, Suite 3500

CITY: Denver

STATE: Colorado

COUNTRY: USA

ZIP: 80203

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08749,699

FILING DATE:

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: Connell, Gary J.

REGISTRATION NUMBER: 32,020

REFERENCE/DOCKET NUMBER: 2618-25-C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 863-9700

TELEFAX: (303) 863-0223

INFORMATION FOR SEQ ID NO: 30:

SEQUENCE CHARACTERISTICS:

LENGTH: 400 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-749-699-30

Alignment Scores:

Pred. No.: 4,4e-39 Length: 400

Score: 432.50 Matches: 101

Percent Similarity: 53.85% Conservative: 46

Best Local Similarity: 37.00% Mismatches: 91

Query Match: 25.93% Indels: 35

DB: 11 Gaps: 11

US-09-735-713A-1 (1-921) x US-08-749-699-30 (1-400)

```
QY 106 AGTTGGGCGAGAGTCTGTTAAGGTACAGCCTTGAATTTTAAACATTTTCACTGCGC 165
Db 148 ThrCysGlyGlu-----LeuTyrThrArgSerAsnArg 158
QY 166 ATTCTGGAGGACCAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAACAA 225
Db 159 IleValGlyGlyHisSerThrGlyPheGlySerHisProtrpGlnAlaLeuLys 178
QY 226 -----AGGCAGGAAGCATATTGTGGAGGAGCATCTCTCACCACAGTGGTG 273
Db 179 SerGlyPheLeuSerLysLysLeuSerCysGlyGlyAlaLeuValSerAspArgTrpVal 198
QY 274 ATCAGCGCGCTCAGTCATGTCGAATGCAACAGAACATTGTGCTACTTTGAATGTACTGCT 333
Db 199 IleThrAlaAlaHisCysValAlaThr---ThrProAsnSerAsnLeuLysValArgLeu 217
QY 334 GGAGATATGACTATAGCCAGACAGACCCCA-----GGAGAGCAAACTCTCATTATTGAA 387
Db 218 GlyGluTrpAspValArgAspHisAspGluArgLeuAsnHisGluLufyAlaIleGlu 237
QY 388 ACTGTCATCATCATACATCCACATTTCTCCACCAAGAAACATGAGCATAT-----GATATT 441
```

```
Db 238 ArgLysGluValHisProSerTyrSer-----ProThrAspPheArgAsnAspVal 254
QY 442 GCCCTTTTGAAGATGGCTGGAGCCTTCCAAATTGGCCACTTTGTGGGCCCATATGCTT 501
Db 255 AlaLeuValLysLeuAspArgThrValIlePheLysGlnHisIleLeuProValCysLeu 274
QY 502 CCAGAGCTGCGGGAGCAATTTGAGCGTGTATTTATTTGTACAACTGCAGGCTGGGGCCGC 561
Db 275 ProHis---LysGlnMetLysLeuAlaGlyLysMetAlaThrValAlaGlyTrpGlyArg 293
QY 562 TTAAGTGAAGT---GGCGTCTCTCACAAGCTGTGCAGGAAGTCAATCTCCCTATTTTG 618
Db 294 ThrArgHisGlyGlnSerThrValProAlaValLeuGlnGluValAspValGluValIle 313
QY 619 ACCTGGGAAGAGTGTGTGGCAGCTCTGTTAACTA---AAGAGGCCCATCAGTGGGAAG 675
Db 314 ProAsnGluArgCysGlnArgTrpPheArgAlaAlaGlyArgArgGluThrIleHisAsp 333
QY 676 ACCTTTCTTGCACAGGTTTTCCTGATGGAGGAGAGACGATGTACGGGAGATTTCAGGA 735
Db 334 ValPheLeuCysAlaGlyTyrLysGluGlyGlyArgaspSerCysGlnGlyAspSerGly 353
QY 736 GGTTCACATGTCGCCGAATAAGAAAGGGCCCTGGACTCTGGCTGGTGTGACTTCCTGG 795
Db 354 GlyProLeuIleMetGln---IleGluGlyArgArgThrLeuValGlyLeuValSerTrp 372
QY 796 GGTTCGGCTGTGTGCGAGGCTGGAGAAACAATGTGAGGAAAGTGCATCAGGATCCCT 855
Db 373 GlyIleGlyCysGlyArg-----GluHisLeuPro 382
QY 856 GGGATCTTCACAGACATTAGTAAAGTGTCTTCCTGGATC 894
Db 383 GlyValTyrThrAsnIleGlnLysPheIleProTrpIle 395
```

Search completed: March 25, 2003, 07:01:00  
Job time : 20 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model  
Run on: March 25, 2003, 06:31:51 ; Search time 42.5 Seconds  
(without alignments)  
4166.579 Million cell updates/sec

Title: US-09-735-713A-1  
Perfect score: 1668  
Sequence: 1 atgagtctcaaatgcttat.....acatccaaactggtactaa 921

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283224 seqs, 96134422 residues  
Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_plus\_n2p.model -DEV=xlh  
-Q/cn2\_1/USPO\_SPOOL/US09735713/runat\_18032003\_124354\_23139/app\_query.fasta\_1.1095  
-DB=PIR\_73 -OFFMT=fastan -SUFFIX=ipr -MINMATCH=0.1 -LOOPECL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09735713 @CGN.1.1.26. @runat\_18032003\_124354\_23139 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEROUT=120  
-WARN\_TIMEROUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR\_73:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	690.5	41.4	oviductin (EC 3.4.21.1)
2	626	37.5	polyprotein - Afri
3	486.5	29.2	serine proteinase
4	466	27.9	plasma kallikrein
5	464	27.8	plasma kallikrein
6	446	26.7	plasma kallikrein
7	442.5	26.5	enteropeptidase (E
8	441	26.4	proctasin (EC 3.4.21.1)
9	440.5	26.4	enteropeptidase (E
10	440	26.4	hepsin (EC 3.4.21.1)
11	438	26.3	hepsin (EC 3.4.21.1)
12	433	26.0	coagulation factor
13	431.5	25.9	enteropeptidase (E
14	421	25.2	membrane-bound arg

15	421	25.2	855	2	JC7775	membrane type-seri
16	420.5	25.2	1047	2	A55617	masquerade precurs
17	419.5	25.1	237	1	TRCY1	trypsin (EC 3.4.21
18	419	25.1	264	2	I38136	chymotrypsin-like
19	415.5	24.9	275	2	S40007	trypsin (EC 3.4.21
20	415	24.9	625	1	KFH01	coagulation factor
21	414.5	24.9	615	1	KFH12	chymotrypsin (EC 3
22	413.5	24.8	263	2	A21195	chymotrypsin (EC 3
23	409	24.5	263	2	A31299	chymotrypsin (EC 3
24	408.5	24.5	415	1	A34170	acrosin (EC 3.4.21
25	408	24.5	275	2	S40005	trypsin (EC 3.4.21
26	407.5	24.4	274	2	S35339	trypsin (EC 3.4.21
27	406	24.3	254	1	TRW33Y	trypsin-like prote
28	405.5	24.3	603	2	S28941	coagulation factor
29	405	24.3	265	2	T15451	hypothetical prote
30	404	24.2	407	1	KFB07	coagulation factor
31	402.5	24.1	275	2	C35863	trypsin (EC 3.4.2
32	402	24.1	245	1	KYBOB	chymotrypsin (EC 3
33	400	24.0	263	1	KYRTB	chymotrypsin (EC 3
34	400	24.0	275	2	A32410	trypsin (EC 3.4.2
35	399.5	24.0	248	2	S55066	trypsin (EC 3.4.21
36	397.5	23.8	248	2	S55067	trypsin (EC 3.4.21
37	397.5	23.8	461	1	JX0210	protein C (activat
38	397.5	23.8	812	1	PLBO	plasmin (EC 3.4.21
39	396.5	23.8	461	1	KFHU	coagulation factor
40	396	23.7	267	2	S40006	trypsin (EC 3.4.21
41	395	23.7	229	1	TRDFS	trypsin (EC 3.4.21
42	395	23.7	1113	2	JR0315	low-density lipopr
43	394	23.6	263	2	S47537	chymotrypsin (EC 3
44	393.5	23.6	275	2	B35863	trypsin (EC 3.4.2
45	392.5	23.5	269	2	B32410	mastocytoma protei

ALIGNMENTS

RESULT 1

T30338

Oviductin (EC 3.4.21.1) - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C>Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 01-Dec-2000

C:Accession: T30338; A00422

R:Lindsay, L.L.; Wieduwilt, M.J.; Hedrick, J.L.

Biol. Reprod. 60, 989-995, 1999

A:Title: Oviductin, the Xenopus laevis oviductal protease that processes egg envelope composed of two protease and several CUB domains.

A:Reference number: Z20830; MUID:99184825; PMID:10084976

A:Accession: T30338

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-1004 <LIN>

A:Cross-references: EMBL:U81291; NID:gl754713; PID:gl754714; PIDN:AAB53972.1

R:Hardy, D.M.; Hedrick, J.L.

Biochemistry 31, 4466-4472, 1992

A:Title: Oviductin purification and properties of the oviductal protease that process

A:Reference number: A40242; MUID:92256375; PMID:1581303

A:Accession: A40242

A:Molecule type: protein

A:Residues: 46-73 <HAR>

C:Comment: This protease is found in oviductal secretory granules and is secreted to

C:Superfamily: trypsin homology

C:Keywords: extracellular protein; hydrolase; serine proteinase

Alignment Scores:			
Pred. No.:	1.86e-55	Length:	1004
Score:	690.50	Matches:	132
Percent Similarity:	60.33%	Conservative:	49
Best Local Similarity:	44.00%	Mismatches:	110
Query Match:	41.40%	Indels:	9
DB:	2	Gaps:	4

US-09-735-713A-1 (1-921) x T30338 (1-1004)

QY 22 ACAGGAACAAGCTGATTTTACTAGGAATAGTCTTTTGTGAACRAGGTAATCTGCA 81

```
Db      3  ThrArgAsn-----LeuLeuLeuGlySerIleLeuLeuSerLeuAlaValLysGly 19
QY      82  RCTCTTCGCTCCCAAGAGCTCCAGTGTGGGAGAGT---CTGGTTAAGGTACAGCCT 138
Db      20  AspProGlyProHisArgGlyAlaArgCysGlyValSerProLeuGlySerAlaThrGlu 39
QY      139  TGAATATTATTAACATTTTCAGTCCGATTTCTGGAGGAGCCAGCAAGTGGAGAGGCTTCC 198
Db      40  LeuAsnTyr-----LeuSerArgIleValGlyArgGlnSerLysGlyGln 56
QY      199  TATCCCTGGCAGGTATCTCTGAACAAAGGAGAGCATATTTGTGGAGAGCATCGTC 258
Db      57  HisProIrpThrValSerLeuLysArgAsnGlyLysHisPheCysGlyThrLeuVal 76
QY      259  TCACCACAGTGGGTATCATCGCGGCTCACTGATTCGAAACAGAACATTTGTGCTACT 318
Db      77  SerHisCysHisValLeuThrAlaAlaHisCysLeuLeuAspArgAsnValLysLeuTyr 96
QY      319  TTGAATGTTACTGCTGGAGAGTATGACTTAAAGCCAGACAGACCCAGAGAGCAAACTCTC 378
Db      97  MetArgValTyrIleGlyGluTyrAspGlnIleLeuLysGluThrGluGlnMetPhe 116
QY      379  ACTATTGAACACTGTATCATACATCCACATTTCTCCACCAAGAAACCAATGGATGAT 438
Db      117  ArgValIleGluIlePheLysHisProAsnPheAsnGlnSerGlnProMetAsnTyrAsp 136
QY      439  ATTGCCCTTTTGAAGATGGCTGAGCCTTCCAAATTTGGCCACTTTTGGGGCCCATATGT 498
Db      137  ValAlaValLeuLeuLeuAspGlySerValThrPheAspGluAsnIleGlnProAlaCys 156
QY      499  CTTCCAGAGCTGCGGGAGCAATTTGAGGCTGGTATTATTGTACAACCTGCAGGCTGGGC 558
Db      157  LeuProAsnProAspValPheGluProGlyAspLeuCysValThrLeuGlyTrpGly 176
QY      559  CGTTTAACGAAGTGGCTCTCTCACAGTCTTCAGAGAAAGTAATGCCTATTTTG 618
Db      177  HisLeuThrGluAsnGlyIleLeuProValValLeuGlnGluValTyrLeuProIleVal 196
QY      619  ACCTGGGAAGAGTGTGGCAGCTCTGTTAAACACTAAAGAGGCCCATCAGTGGGAAGACC 678
Db      197  AspLeuSerSerCysLeuHisValMetSerAlaLeuLysGlyThrValValSerTyr 216
QY      679  TTTCTTTGCACAGGTTTCTGATGAGGAGAGACGCATGTCCAGGAGATTCAGGAGGT 738
Db      217  IleValCysAlaGlyPheProGluGlyGlyLysAspAlaCysGlnGlyAspSerGlyGly 236
QY      739  TCACATGTCGCCGGAATAAGAAGGGCCTGGACTCTGGCTGGTGTGACTTCTCGGGT 798
Db      237  ProLeuLeuCysGlnArgArgHisGlySerTrpValLeuHisGlyLeuThrSerTrpGly 256
QY      799  TTGGGCTGTGCTGAGGCTGGAGAACAAATGTG-----AGGAAAGTGATCAAGGATCC 852
Db      257  MetGlyCysGlyArgSerTrpLysAsnAsnValPheLeuProHisAsnArgLysGlySer 276
QY      853  CTGGGATCTTCACAGACATTAGTAAGTCTTTCTTCGATCCACCAAGAACATCCAAACT 912
Db      277  ProGlyIlePheThrAspIleGlnLysLeuLeuGlyTrpValSerSerGlnLeuAsnThr 296

RESULT 2
T30337
polyprotein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30337
R:Yang, J.-C.; Lindsey, L.L.; Hedrick, J.L.
submitted to the EMBL Data Library, March 1998
A:Description: cDNA cloning of ovochymase, a chymotrypsin-like protease released from xe
A:Reference number: Z20829
A:Accession: T30337
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1524 <YAN>
```

```
A:Cross-references: EMBL:U81290; NID:g2981640; PID:g2981641; PIDN:AAC24717.1
Alignment Scores:
Pred. No.:      1.8e-49      Length:      1524
Score:          626.00      Matches:     128
Percent Similarity: 57.89%      Conservative: 37
Best Local Similarity: 44.91%      Mismatches:  110
Query Match:     37.53%      Indels:      10
DB:              2          Gaps:          4

US-09-735-713a-1 (1-921) x T30337 (1-1524)
QY      70  GGTAAATCTGCARCTCTTTTCGCTCCCAAGCTCCAGTGTGGGAGAGTCTGGTTAAG 129
Db      32  GlyGluThrAlaGluLeuLysCysGlyThrArgProGluIleGlyAsp----- 47
QY      130  GTACAGCCTTGGAAATTTATTTTAACTATTTTCAGTCGCATTTCTTGAGGAGAACCCAGTGGAG 189
Db      48  --GluProAspLeuGluPheThr-----SerArgIleValGlyGlyGlyAspAlaAla 64
QY      190  AAGGTTTCCTATCCCTGGCAGGATCTCTGAAACAAGGAGCAGAGCATATTTGTGGAGGA 249
Db      65  ValGlyGlyGlnProIrpThrValSerLeuLysLeuAsnGluArgHisIleCysGlyGly 84
QY      250  AGCATCTCTCACACAGTGGGTGATCAGCGCGCTCACTGCATT-----GCAAAACAGA 303
Db      85  SerIleValArgLysAspMetValValThrAlaAlaHisCysValTyrProValThrGlu 104
QY      304  ACATTGTGCTACTTGTGAATGTTACTGTGTGAGAGATGATGACATTAAGCCAGACAGACCCA 363
Db      105  IleLysValSerHisMetThrValIleValGlyGlyTyrAspGlnGlnValMetAspSer 124
QY      364  GGAGAGCAACCTCACTATTGAACCTGTCATCATACATCCACATTTCTCCACCAAGAAA 423
Db      125  GlnGluGlnSerIleProValSerHisIleGluProHisProAsnTyrArgGlyAspGly 144
QY      424  CCAATGACATGATGATATTCCTTTTGAAGATGGCTGGAGCCCTTCCAATTTTGGCCACTTT 483
Db      145  AsnMetGlyTyrAspIleAlaLeuValPheLeuSerLysProIleIlePheGlySerGln 164
QY      484  GTGGGGCCCATATGCTTTCAGAGCTGCGGGAGCAGCATTTTTCAGCTGGTATTTATTGTACA 543
Db      165  ValGlnProIleCysLeuProGlnValGlyGlyLysIleGluAlaGlyThrLeuCysVal 184
QY      544  ACTGAGGCTGGGGCCGCTTAAGTGAAGTGGCTCTCTCAACAGCTCTTGCAGGAAGTG 603
Db      185  SerSerGlyTrpGlyArgLeuGluGluAsnGlyAspLeuSerProValLeuGlnVal 204
QY      604  AATCTGCCCTATTATTGACCTGGGAAGAGTGTGTGGAGAGTCTGTAAACACTAAAGAGGCC 663
Db      205  LysLeuProValValAspAsnGlyThrCysHisAlaValLeuGluProIleGlyHisPro 224
QY      664  ATCAGTGGGAAGACCTTTCTTTCACAGAGGTTTTCCTGTATGGAGGGAGACGCGATGTCAG 723
Db      225  ValLeuAspAspThrMetLeuCysAlaGlyPheProGluGlyGlyMetAspAlaCysGln 244
QY      724  GGAGATTACAGAGGTTTCACTCATGTGCGGAATAAAGAGGCGCTGGACTCTGGCTGGT 783
Db      245  GlyAspSerGlyGlyProPheValCysArgArgArgSerGlyValIrpPheLeuAlaGly 264
QY      784  GTGACTTCTCGGGTTTGGGCTGTGGTCGAGGCTGG---AGAAACAATGTGAGGAAAGT 840
Db      265  CysValSerTrpGlyLeuGlyCysGlyArgSerTrpGlyAlaLysGlnIleIleArgSer 284
QY      841  GATCAAGGATCCCTGGGATCTTCACAGACATTAAGTAAAGTCTTCTCGGATCCAGGAA 900
Db      285  GlnSerGlySerProAlaIlePheSerArgValSerSerValLeuAspPheLeuArgPro 304
QY      901  CACATCCAAACTGGT 915
Db      305  ProLysLeuThrGly 309

RESULT 3
```



## Alignment Scores:

Pred. No.: 9,88e-35 Length: 638  
Score: 466.00 Matches: 91  
Percent Similarity: 55.47% Conservative: 51  
Best Local Similarity: 35.55% Mismatches: 96  
Query Match: 27.94% Indels: 18  
DB: 1 Gaps: 5

US-09-735-713a-1 (1-921) x KQHUP (1-638)

```
QY 160 AGTCGATCTTCGGAGAGCAAGTGGAGAGGGTTCATCCCTGGCAGGTATCTCTG 219
Db 389 ThrArgIleValGlyGlyThrAsnSerSerrpGlyGluTrpProfrpGlnValSerLeu 408
QY 220 AAACAAGG-----CAGAAGCATATTGTGGAGGAAGCATCTCTCACACAGTGG 270
Db 409 GlnValLysLeuThralaGlnArgHisLeuCysGlySerLeuIleGlyHisGlnTrp 428
QY 271 GTGATCAGCGCGCTCAGTCATTCGAAACAGAAACATTTGTCTACTTTGAATGTACT 330
Db 429 ValLeuThrAlaAlaHisCysPheAspGlyLeuProLeuGlnAspValTrpArgIleTyr 448
QY 331 GCTGGAGATGACTTAAGTACAGCAGACAGCCAGAGCAAACTCTCACTATTGAAACT 390
Db 449 SerGlyIleLeuAsnLeuSerAspIleThrLysAspThrProPheSerGlnIleLysGlu 468
QY 391 GTCATCATATCCACATTTCTCCACCAAGAACCAATGGACTATGATATTGCCCTTTG 450
Db 469 IleIleIleHisGlnAsnTrpLysValSerGlyGly---AsnHisAspIleAlaLeuIle 487
QY 451 AGATGGCTGGAGCCCTCCAAATTTGGCCACTTTGGGCGCCCATATGCTTCCAGAGCTG 510
Db 488 LysLeuGlnAlaProLeuAsnTrpThrGluPheGlnLysProIleCysLeuProSerLys 507
QY 511 CGGAGCAATTTGAGGCTGTTTATTCTTCAACTGCAGGCTGGGCGCGCTTAAGTAA 570
Db 508 GlyAspThrSerThrIleThrLysCysTrpValThrGlyTrpGlyPheSerLysGlu 527
QY 571 GGTGGCTCCTCTCACAGCTTGTGAGGAAGTGAATCTGCCTTATTGACCTGGGAAGAG 630
Db 528 LysGlyGluIleGlnAsnIleLeuGlnLysValAsnIleProLeuValThrAsnGluGlu 547
QY 631 TGTGTGGAGCTCTGTATACACTAAGAGGCCCATCATGAGGAAGACCTTTCTTGGACA 690
Db 548 CysGlnLysArgTrpGlnAspTrpLys-----IleThrGlnArgMetValCysAla 564
QY 691 GGTTCCTTCATGGAGGAGAGCAGCATCTCAGGAGATTCAGGAGTTCACCTCATGTGC 750
Db 565 GlyTrpLysGlyGlyLysAspAlaCysLysGlyAspSerGlyGlyProLeuValCys 584
QY 751 CGGAATAAGAAAGGCGCTGGACTCTGGCTGTGTCACCTCTCTGGGTTGGCTGTGTGT 810
Db 585 LysHis---AsnGlyMetTrpArgLeuValGlyIleThrSerTrpGlyGlyCysAla 603
QY 811 CGAGGCTGGAGAACAAATGTGAGGAAAGTATCATAGGATCCCTGGGATCTTCACAGAC 870
Db 604 Arg-----ArgGluGlnProGlyValTrpThrLys 613
QY 871 ATTAGTAAGTCTTTCCTGGATCCAGCAACACATCCAACTGGTAAC 918
Db 614 ValAlaGluTrpMetAspTrpIleLeuGluLysThrGlnSerSerAsp 629
```

## RESULT 5

## KOMSP

Plasma kallikrein (EC 3.4.21.34) precursor - mouse

C:Species: Mus musculus (house mouse)

C&gt;Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 18-Jun-1999

C:Accession: A36557

R:Seidah, N.G.; Sawyer, N.; Hamelin, J.; Mion, P.; Beaubien, G.; Brachpapa, L.; Rochemond

DNA Cell Biol. 9, 737-748, 1990

A:Title: Mouse plasma kallikrein: cDNA structure, enzyme characterization, and compariso

A:Reference number: A36557; PMID:91090844; PMID:2264928

A:Accession: A36557

A:Molecule type: mRNA

A:Residues: 1-638 &lt;SEI&gt;

A:Cross-references: GB:MS8588; NID:g200358; PIDN:AAA63393.1; PID:g200359

A:Note: part of this sequence, including the amino ends of both the heavy and light c

C:Comment: This protein, synthesized in the liver, circulates as a noncovalent comple

C:Comment: The zymogen is activated by factor XIIa, which cleaves the molecule into a

are linked by one or more disulfide bonds.

C:Superfamily: coagulation factor XI; trypsin homology

C:Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; in

F:1-19/Domain: signal sequence #status predicted &lt;SIG&gt;

F:20-390/Product: plasma kallikrein heavy chain #status experimental &lt;HCH&gt;

F:20-109/Domain: apple repeat &lt;AP1&gt;

F:110-199/Domain: apple repeat &lt;AP2&gt;

F:200-289/Domain: apple repeat &lt;AP3&gt;

F:291-380/Domain: apple repeat &lt;AP4&gt;

F:391-638/Product: plasma kallikrein light chain #status experimental &lt;LCH&gt;

F:391-621/Domain: trypsin homology &lt;TRY&gt;

F:21-104/Domain: tryptophan 194,137-166,141-147,201-284,227-256,231-237,292-375,318-347,

F:127,215,308,396,494/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:434,483,578/Active site: His, Asp, Ser #status predicted

## Alignment Scores:

Pred. No.: 1.51e-34 Length: 638  
Score: 464.00 Matches: 95  
Percent Similarity: 53.49% Conservative: 43  
Best Local Similarity: 36.82% Mismatches: 102  
Query Match: 27.82% Indels: 18  
DB: 1 Gaps: 6

US-09-735-713a-1 (1-921) x KOMSPL (1-638)

```
QY 154 ATTTTCAGTCGATCTTCGGAGGAAGCCAAAGTGGAGAGGGTTCCTATCCCTGGCAGGTA 213
Db 387 IleAsnAlaArgIleValGlyGlyThrAsnAlaSerLeuGluTrpProfrpGlnVal 406
QY 214 TCTCTGAACAAGG-----CAGAAGCATATTGTGGAGGAAGCATCTCTCACCA 264
Db 407 SerLeuGlnValLysLeuValSerGlnThrHisLeuCysGlySerIleIleGlyArg 426
QY 265 CAGTGGGTGATCAGCGCGCTCAGTCATTCGAAACAGAAACATCTGTCTACTTTGAAT 324
Db 427 GlnTrpValLeuAlaAlaHisCysPheAspGlyIleProTrpProAspValTrpArg 446
QY 325 GTTACTGCTGGAGAGTATGACTTAAGCCAGACAGACCCAGGAGCAAACTCTCACTATT 384
Db 447 IleTyrGlyGlyIleLeuSerLeuSerGluIleThrLysGluThrProSerSerArgIle 466
QY 385 GAACTGTCATCATCATCATCCATTTCTCCACCAAGAACCAATGGACTATGATATTGCC 444
Db 467 LysGluLeuIleIleHisGlnGluTrpLysValSerGluGly---AsnTrpAspIleAla 485
QY 445 CTTTTCAGATGGCTGGAGCCCTTCCAAATTTGGCCACTTTGTGGGCCCCATATCTCTTCA 504
Db 486 LeuIleLysLeuGlnThrProLeuAsnTrpThrGluPheGlnLysProIleCysLeuPro 505
QY 505 GAGCTGGGAGCAATTTGAGGCTGTTTATTGTTTACAACTGCAGGCTGGGCGCGCTTA 564
Db 506 SerLysAlaAspThrAsnThrIleThrAsnCysTrpValThrGlyTrpGlyTrpThr 525
QY 565 ACTGAAGGTGGCTCCTCTCACAGTCTTTCAGGAGAGTGAATCTGCTATTTTGGACCTGG 624
Db 526 LysGluGlnGlyGluThrGlnAsnIleLeuGlnLysAlaThrIleProLeuValProAsn 545
QY 625 GAAGAGTGTGGCAGCTCTGTTTAACTAAAGAGCCCCATCAGTGGGAGACCTTTCTT 584
Db 546 GluGluCysGlnLys-----LysTrpArgAspTrpValIleAsnLysGlnMetIle 562
QY 685 TGCACAGGTTTTCCTGATGGAGGAGAGACGATGTCCAGGAGATTCAGGAGGTTTCACATC 744
Db 563 CysAlaGlyTrpLysGluGlyGlyThrAspAlaCysLysGlyAspSerGlyGlyProLeu 582
QY 745 ATGTCCCGGAATPAGAAAGGCGCTTGGACTCTGGCTGTGTGACTTCTCTGGGTTGGGC 804
Db 745 ATGTCCCGGAATPAGAAAGGCGCTTGGACTCTGGCTGTGTGACTTCTCTGGGTTGGGC 804
```

Pred. No.:	6.97e-33	Length:	638
Score:	446.00	Matches:	91
Percent Similarity:	53.12%	Conservative:	45
Best Local Similarity:	35.55%	Mismatches:	102
Query Match:	26.74%	Indels:	18
DB:	1	Gaps:	5
US-09-735-713A-1 (1-921) x KQRTPL (1-638)			
QY	154	ATTTTCAGTCGCATCTTGGAGAACCCAAAGTGAGAACGGTTCCTATCCCTGGCAGGTA	213
DB	387	IIeasAlaarglleValGlyGlyThrAsnSerSerLeuGlyGluTrpProTrpGlnVal	406
QY	214	TCCTGAAACAAAGG-----CAGAAGCATATTGTGGAGGAAGCATCGTCTCACCA	264
DB	407	SerLeuGlnValLysLeuValSerGlnAsnHisMetCysGlySerIleIleGlyArg	426
QY	265	CAGTGGTGATACGGGGCTCACTGCATTGCAACAGAAACATTGTGTCTACTTTTGAAT	324
DB	427	GlnTrpIleLeuThrAlaAlaHisCysPheAspGlyIleProTrpProAspValTrpArg	446
QY	325	GTACTGCTGGAGAGTATGACTTAAGCCACAGACAGCCAGGAGAGCAAACTCTCACATAT	384
DB	447	IleTrpGlyGlyIleLeuAsnLeuSerGluIleThrAsnLysThrProPheSerIle	466
QY	385	GAAACTGTCATATACATCCACATTTCTCCACCAAGAACCAATGACATATGATATGGC	444
DB	467	LysGluLeuIleIleHisGlnLysTrpLysMetSerGluGly---SerTrpAspIleAla	485
QY	445	CTTTTGAAGATGCTCGAGCTTCCCAATTGGCCACTTGTGGGGCCCATATGTCTTCCA	504
DB	486	LeuIleLysLeuGlnThrProLeuAsnTrpThrGluPheGlnLysProIleCysLeuPro	505
QY	505	GAGTCGGGAGCAATTTGAGCTGGTTATTATTGTACAACTCCAGCTGGGGCCGCTTA	564
DB	506	SerLysAlaAspThrAsnThrIleTrpThrAsnCysTrpValThrGlyTrpGlyTrpThr	525
QY	565	ACTGAAGTGGCGCTCTCTCACAAAGTCTTGCAGGAAGTGAATCTGCTATTTTGGACCTGG	624
DB	526	LysGluArgGlyGluThrGlnAsnIleLeuGlnLysAlaThrIleProLeuValProAsn	545
QY	625	GAGAGTGTCTGGCAGCTCTGTTAAACACTAAAGAGGCCCATCAGTGGGAAGACCTTTCTT	684
DB	546	GluGluCysGlnLys-----LysTrpArgAspTrpValIleThrLysGlnMetIle	562
QY	685	TGCACAGGTTTTCTGATGGAGGAGAGAGCCATGTTCAGGAGATTCAGGAGGTTTCACTC	744
DB	563	CysAlaGlyTrpLysGluGlyGlyIleAspAlaCysLysGlyAspSerGlyGlyProLeu	582
QY	745	ATGTGCCGGAATAAGAAAGGGCGCTGGACCTCGCTGGTGGTGTGACTTCTCGGGGTTGGGC	804
DB	583	ValCysLysHis---SerGlyArgTrpGlnLeuValGlyIleThrSerTrpGlyGluGly	601
QY	805	TGTGTGCGAGGCTGGAGAAACAATGTGAGSAAAAGTATCAAGGATCCCTCGGATCTTC	864
DB	602	CysAlaArg-----LysGluGlnProGlyValTrp	611
QY	865	ACAGACATTAAGTAAAGTCTTCTCGATCCAGCAACACATCCAAACT	912
DB	612	ThrLysValAlaGluTrpIleAspTrpIleLeuGluLysIleGlnSer	627
RESULT 7			
A43090			
enteropeptidase (EC 3.4.21.9) precursor - bovine			
N:Alternate names: enterokinase			
C:Species: Bos primigenius taurus (cattle)			
C:date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999			
C:Accession: A43090; A48874; A61436			
R:Kitamoto, Y.; Yuan, X.; Wu, Q.; McCourt, D.W.; Sadlier, J.E.			
Proc. Natl. Acad. Sci. U.S.A. 91, 7588-7592, 1994			
A:title: Enterokinase, the initiator of intestinal digestion, is a mosaic			
A:reference number: A43090; MUID:94329561; PMID:8052624			
A:Accession: A43090			

A;Title: Enterokinase, the initiator of intestinal digestion,  
A;Reference number: A43090; MUID:94329561; PMID:8052624

859 LeuHisMetAlaSerAsnLeuThrSerProGlnIleGluThrArgLeuIleAspGlnIle 878  
 394 ATCATACATCCACATTTCCTCCACCAAGAAACCAATGACATATGATATATGCTCTTTTGAAG 453  
 879 ValIleasnProHisTyrAsn---LysArgArgLysAsnAsnAlaMetMethHis 897  
 454 ATGGCTGGAGCCTTCCCAATTTGGCCACTTGTGGGGGCCATATGCTCTCCAGAGCTGCGG 513  
 898 LeuGluMetLysValAsnTyrThrAspTyrIleGlnProIleCysLeuProGluGluAsn 917  
 514 GAGCAATTGAGGCTGGTTTATTTGTTACAACTGCAGGCTGGGCCGCTTAACTCAAGGT 573  
 918 GlnValPheProProGlyArgIleCysSerIleAlaGlyTrpGlyAlaLeuIleTyrGln 937  
 574 GCGTCTCTCACAAAGTTCGAGGAAGTGAATGCTGCTATTTTGACCTGGGAAGAGTGT 633  
 938 GlySerThrAlaAspValLeuGlnGluAlaAspValProLeuLeuSerAsnGluLysCys 957  
 634 GTGCGAGCTCTG-----TTAACACTAAAGAGGCCCATCAGTGGGAAGACCTTT 681  
 958 GlnGlnMetProGluTyrAsnIleThr-----GluAsnMet 970  
 682 CTTTGCACAGGTTTCTGATGGAGGAGAGACGCATGTCCAGGAGATTCAGGAGGTCA 741  
 971 ValCysAlaGlyTyrGluAlaGlyGlyValAspSerCysGlnGlyAspSerGlyGlyPro 990  
 742 CTGATGCGCGAATAAGAAGGGCTGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 801  
 991 LeuMetCysGln---GluAsnAsnArgTrpLeuLeuAlaGlyValThrSerPheGlyTyr 1009  
 802 GGCTGTGCTGAGGCTGGAGCAACAATGTGAGGAAAGTGTATCAAGGATCCCTCGGGATC 861  
 1010 GlnCysAlaLeuProAsnArg-----ProGlyVal 1019  
 862 TTCACAGACATTAGTAAAGTGCTTCTCGATCCACGAACACATC 906  
 1020 TyrAlaArgValProArgPheThrGluTrpIleGlnSerPheLeu 1034  
 RESULT 8  
 A57014  
 prostatic (EC 3.4.21.-) precursor - human  
 C:Species: Homo sapiens (man)  
 C:Date: 24-May-1996 #sequence\_revision 24-May-1996 #text\_change 18-Jun-1999  
 C:Accession: A57014; A54866  
 R:Yu, J.X.; Chao, L.; Chao, J.  
 J. Biol. Chem. 270, 13483-13489, 1995  
 A:Title: Molecular cloning, tissue-specific expression, and cellular localization of  
 A:Reference number: A57014; MUID:95286644; PMID:7768952  
 A:Accession: A57014  
 A:Status: translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-343 <RES>  
 A:Cross-references: GB:I41351; NID:g862304; PIDN:AAC41759.1; PID:g862305  
 A:Experimental source: prostate  
 A:Note: parts of this sequence were determined by protein sequencing  
 R:Yu, J.X.; Chao, L.; Chao, J.  
 J. Biol. Chem. 269, 18843-18848, 1994  
 A:Title: Prostatic is a novel human serine proteinase from seminal fluid. Purification  
 A:Reference number: A54866; MUID:94308140; PMID:8034638  
 A:Accession: A54866  
 A:Molecule type: protein  
 A:Residues: 45-64 <YUA>  
 C:Genetics:  
 A:Gene: GDB:PRSS8  
 A:Cross-references: GDB:676446; OMIM:600823  
 A:Map position: 16p11.2-16p11.2  
 C:Superfamily: prostatic; trypsin homology  
 C:Keywords: glycoprotein; hydrolase; proteinase; transmembrane protein  
 F:1-32/Domain: signal sequence #status predicted <SIG>  
 F:33-44/45-343/product: prostatic #status predicted <MAT>  
 F:45-343/Domain: prostatic light chain #status predicted <CHL>  
 F:45-281/Domain: trypsin heavy chain #status predicted <CHH>  
 F:45-281/Domain: trypsin homology <TRY>

F:323-341/Domain: transmembrane #status predicted <TM>1>  
F:37-154,70-86,168-244,201-223,234-262/Disulfide bonds: #status predicted  
F:85,134,238/Active site: His, Asp, Ser #status predicted  
F:159/Binding site: carbohydrate (Asn) (covalent) #status experimental

Alignment Scores:  
Pred. No.: 1.86e-32 Length: 343  
Score: 441.00 Matches: 100  
Percent Similarity: 49.16% Conservative: 46  
Best Local Similarity: 33.67% Mismatches: 117  
Query Match: 26.44% Indels: 34  
DB: 1 Gaps: 10

US-09-735-713A-1 (1-921) x A57014 (1-343)

QY 34 CTGATTTTACTAGGATAGTCTTTTGAACRAGTAATCGCARTCTTTCGCTC 93  
Db 17 ILeuLeuTyrLeuGlyLeuLeuArgSerGlyThrGlyAlaGluGlyAla----- 33  
QY 94 CCCAAAGCTCCAGTGTGTGGCAGAGTCTGTTAAGGTACAGCCTTGGAAATTTTAAAC 153  
Db 34 ---GluAlaPro---CysGly-----ValAlaProGln----- 42

QY 154 ATTTTCATCGCATCTTGGAGGAGCCAAAGTGGAGAGGTCTCTATCTCCGCGAGGTA 213  
Db 43 -----AlaArgIleThrGlyGlySerSerAlaValAlaGlyGlnTrpProTrpGlnVal 60

QY 214 TCTCTGAACAAAGCCAGAACATATTGTGGAGGAACATCGTCTCACCAGTGGGTG 273  
Db 61 SerIleThrTyrGluGlyValHisValCysGlyGlySerLeuValSerGluGlnTrpVal 80

QY 274 ATCAGCGCGCTCACTGCTATTCACCAAGAACCAATGTGTCTACTTTGAATGTACTGCT 333  
Db 81 LeuSerAlaAlaHisCysPheProSerGluHisHisLysGluAlaTyrGluValLysLeu 100

QY 334 GGAGAGTATGACTTAAGCCAGACAGACCCAGAGAGCAACTCTCATATTGAACACTGTC 393  
Db 101 GlyAlaHisGlnLeuAspSerTyrSerGluAspAlaLysValSerThrLeuLysAspIle 120

QY 394 ATCATACATCACATTTCTCCACCAAGAACCAATGGACTATGATATTGCCCTTTTGAAG 453  
Db 121 IleProHisProSerTyr---LeuGlnGlySerGlnGlyAspIleAlaLeuLeuGln 139

QY 454 ATGGCTGGAGCTTCCAAATTTGGCCACTTTGTGGGCGCCATGTCCTCCAGAGTGGCG 513  
Db 140 LeuSerArgProIleThrPheSerArgTyrIleArgProIleCysLeuProAlaAlaAsn 159

QY 514 GAGCAATTTGAGCTGTTTATTGTACAACTGCAGCGCTGGCGCGCTTAACTGAAGGT 573  
Db 160 AlaSerPheProAsnGlyLeuHisCysThrValThrGlyTyrGlyHisValAlaProSer 179

QY 574 GCGGTCTC-----TCAAGAGTCTTGCAGGAAGTGAATCGCTATTTTGCACCTGGAA 627  
Db 180 ValSerLeuLeuThrProLysProLeuGlnGlnLeuGluValProLeuIleSerArgGlu 199

QY 628 GAGTGTGGCAGCTCTGTTAACTAAAGAGAGGCC-----ATCAGTGGGAG 675  
Db 200 ThrCysAsnCysLeuTyrAsnIleAspAlaLysProGluGluProHisPheValGlnGlu 219

QY 676 ACCTTTCTTTGCACAGTTTCTCGATGGAGGAGACGATGTCAGGAGATTCAGGA 735  
Db 220 AspMetValCysAlaGlyTyrValGluGlyGlyLysAspAlaCysGlnGlyAspSerGly 239

QY 736 GGTTCACATGTCGCGGAATAAGAAAGCGCTGGACTCTGGCTGTGTCATCTCCG 795  
Db 240 GlyProLeuSerCys---ProValGluGlyLeuTrpTyrThrGlyIleValSerTrp 258

QY 796 GGTGTGGCTGTGGTCGAGGCTGGAGAAACAATGTGAGGAAAGTGAATCAAGGATCCCT 855  
Db 259 GlyAspAlaCysGlyAla-----ArgAsnArg-----Pro 268

QY 856 GGGATCTTCACAGACATTAGTAAAGTGTCTTCCTGGATCCAGAACATC 906  
|||||:||||| ||| |||||||

Db 269 GlyValTyrThrLeuAlaSerSerTyrAlaSerTrpIleGlnSerLysVal 285  
RESULT 9  
A53663  
enteropeptidase (EC 3.4.21.9) precursor - pig  
N:Alternate names: enterokinase  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 07-Oct-1994 #sequence\_revision 09-Aug-1996 #text\_change 18-Jun-1999  
C:Accession: A53663  
R:Matsushima, M.; Ichinose, M.; Yahagi, N.; Kakel, N.; Tsukada, S.; Miki, K.; Kuroka  
J. Biol. Chem. 269, 19976-19982, 1994  
A:Title: Structural characterization of porcine enteropeptidase.  
A:Reference number: A53663; MUID:94327548; PMID:8051081  
A:Accession: A53663  
A:Molecule type: mRNA  
A:Residues: 1-1034 <MAT>  
A:Cross-references: GB:D30799; NID:g505122; PIDN:BAA06459.1; PID:g505123  
A:Note: parts of this sequence, including the amino ends of three chains isolated fr  
C:Comment: The mechanism of association with the membrane of the intestinal brush bo  
ated below) or with amino-terminal myristoylation of the heavy chain.  
C:Complex: Mature enteropeptidase is variously reported to contain two (heavy and l  
ed by a disulfide bond. Possibly, conversion from membrane-bound to soluble forms in  
C:Function:  
A:Description: cleaves activation peptide from trypsinogen to produce active trypsin  
A:Pathway: intestinal digestive hydrolase cascade  
C:Superfamily: enteropeptidase; Clr/Cls repeat homology; LDL receptor ligand-binding  
C:Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein; zymog  
F:22-38/Domain: transmembrane #status predicted <TM>  
F:52-117/Product: enteropeptidase mini chain #status predicted <MCH>  
F:118-799/Product: enteropeptidase heavy chain #status predicted <HCH>  
F:199-236/Domain: LDL receptor ligand-binding repeat homology <LDL1>  
F:357-519/Domain: MAM homology <MAM>  
F:541-646/Domain: Clr/Cls repeat homology <Clr>  
F:658-692/Domain: LDL receptor ligand-binding repeat homology <LDL2>  
F:693-798/Domain: scavenger receptor cysteine-rich domain homology #status atypical  
F:800-1034/Product: enteropeptidase light chain #status predicted <LCH>  
F:800-1029/Domain: trypsin homology <TRY>  
F:116,147,170,194,283,343,350,403,455,485,518,549,645,697,701,721,740,761,804,863,90  
F:787-911,925-841,925-992,956-971,982-1010/Disulfide bonds: #status predicted  
F:840,891,986/Active site: His, Asp, Ser #status predicted

Alignment Scores:  
Pred. No.: 2.4e-32 Length: 1034  
Score: 440.50 Matches: 101  
Percent Similarity: 48.95% Conservative: 62  
Best Local Similarity: 30.33% Mismatches: 93  
Query Match: 26.41% Indels: 77  
DB: 1 Gaps: 11

US-09-735-713A-1 (1-921) x A53663 (1-1034)

QY 43 CTACTAGGAATA-----GTCTTTTGAACRAGGTAATCTGCA 81  
|||||:|||||:||||| ||||| |||  
Db 733 LeuLeuGlyLeuGlyThrGlyAsnSerSerMetProPhePheSerSerGlyGlyPro 752

QY 82 RCTCTTTTCGCTCCCAAGAGCTCCC----- 105  
||| ||| |||||  
Db 753 PheValLysLeuAsnThrAlaProAsnGlySerLeuIleLeuThrAlaSerGluGlnCys 772

QY 106 -----AGGTACAGCCTTGAATATTATTTAACATTTTCAGTCGATCTTGGAGAGCC 180  
|||||:|||||:||||| |||||:|||||:||||| |||  
Db 773 PheGluAspSerLeuIleLeuLeuGlnCysAsnHisLysSerCysGlyLysGlnVal 792

QY 127 -----AGGTGGAAGGTTCTCTATCCCTGGCAGGTATCTCTGAAACAAGAGCAGAGCATATT 240  
|||||:|||||:||||| |||||:|||||:||||| |||  
Db 181 CAAGTGGAGAGGTTCTCTATCCCTGGCAGGTATCTCTGAAACAAGAGCAGAGCATATT 240

QY 805 AspSerArgGluGlyAlaTrpProTrpValValAlaLeuTyrTyrAsnGlyGlnLeuLeu 824  
|||||:|||||:||||| |||||:|||||:||||| |||

QY 241 TGTGAGGAGAGCATCGTCTCACCAGGTGGGTGATCAGCGGGCTCACTGTCATTCGCAAC 300  
|||||:|||||:||||| |||||:|||||:||||| |||



A:Reference number: S33777; MUID:93305733; PMID:8318546

A:Accession: S33777

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-416 <FAR>

A:Cross-references: EMBL:X70900; NID:g57928; PIDN:CAA50256.1; PID:g57929

C:Superfamily: hepsin; trypsin homology

C:Keywords: hydrolase; liver; serine proteinase; transmembrane protein

F:22-44/Domain: transmembrane #status predicted <TMN>

F:162-399/Domain: trypsin homology <TRY>

F:187-203,290-358,321-337,348-380/Disulfide bonds: #status predicted

F:202,256,352/Active site: His, Asp, Ser #status predicted

#### Alignment Scores:

Pred. No.:	3,62e-32	Length:	416
Score:	438.00	Matches:	99
Percent Similarity:	49.47%	Conservative:	41
Best Local Similarity:	34.98%	Mismatches:	99
Query Match:	26.26%	Indels:	44
DB:	1	Gaps:	10

US-09-735-713A-1 (1-921) x S33777 (1-416)

```
QY 109 TGTGGCAGAGCTCGTTAAGTACAGCTTGGAAATTTTAAACATTTTCAGTCGCATT 168
|||||
Db 152 CysGlyArgArgLysLeuProVal-----AspArgile 162

QY 169 CTTGAGGAGCCAAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAAGG 228
|||||
Db 163 valGlyGlnAspSerSerLeuGlyArgTrpTrpGlnValSerLeuArgTyrAsp 182

QY 229 CAGAAGCATATTTGTGGAGGAAGCATGCTCTACACAGAGTGGGTGATCAGCGCGCTCAC 288
|||||
Db 183 GlyThrHisLeuCysGlySerLeuSerGlyAspTrpValLeuThrAlaAlaHis 202

QY 289 TGCATTGCAACAGAAC---ATTGTGCTACTTTGAATGTACTGTCTGGAGATGAC 345
|||||
Db 203 CysPheProGluArgAsnArgValLeuSerArgTrpArgValPheAlaGlyAla----- 220

QY 346 TTAAGCAGACAGACCAGGAGAGCAAACTCTCACTATTGAACTGTGCATCATCATCCA 405
|||||
Db 221 ValAlaArgThrSerProHisAlaValGlnLeuGlyValGlnAlaValIleThrHisGly 240

QY 406 CATTTC-----TCCACCAAGAAACCAATGGACTATGATATTGCTCCTTTTG 450
|||||
Db 241 GlyTyrLeuProPheArgAspProThrIleAspGluAsnSerAsnAspIleAlaLeuVal 260

QY 451 AAGATGGCTGGAGCCTTCCCAATTTGGCCACTTTGTGGGGCCCATATGCTTCCAGAGCTG 510
|||||
Db 261 HisLeuSerSerLeuProLeuThrGluTyrIleGlnProValCysLeuProAlaAla 280

QY 511 CGGGAGCAATTTGAGGCTGTTTATTTGTACAACTGCAGGCTGGGGCCGCTTAACGTAA 570
|||||
Db 281 GlyGlnAlaLeuValAspGlyLysValCysThrValThrGlyTrpGlyAsnThrGlnPhe 300

QY 571 GGTGGCTCTCTCACAGTCTTCAGGAAGTGAATCTGCCTATTATTTGACCTGGGAAGAG 630
|||||
Db 301 TyrGlyGlnGlnAlaValLeuGlnGlnAlaArgValProIleIleSerAsnGluVal 320

QY 631 TGTGTGCAGCTGTGTTAACTAAAGAGGCC-----ATCAGTGGG 672
|||||
Db 321 Cys-----AsnSerProAspPheTyrGlyAsnGlnIleLysPro 333

QY 673 AAGACCTTCTTCACAGGTTTTCCTGATGGAGGAGAGCCGATGTCAGGGAGATTCA 732
|||||
Db 334 LysMetPhe---CysAlaGlyTyrProGluGlyIleAspAlaCysGlnGlyAspSer 352

QY 733 GGAGGTTCACTCATGTGCCGAATAAG-----AAAGGGCCTGCACCTCGCTGGT 783
|||||
Db 353 GlyGlyHisPheValCysGluAspArgIleSerGlyThrSerArgTrpArgLeuCysGly 372

QY 784 GTGACTTCTGGGTTGGGCTGTGGTCGAGGCTGGAGAAACAATGTGAGAAAGTGAT 843
|||||
```

```
Db 373 IleValSerTrpGlyThrGlyCysAlaLeu-----AlaArgLys----- 385
QY 844 CAAGGATCCCTGGATCTTCACAGACATTAGTAAAGTGCTTCTCTGATCCAGCAACAC 903
|||||
Db 386 -----ProGlyValTyrThrLysValIleAspPheArgGluTrpIlePheGlnAla 402
QY 904 ATCCAAACT 912
|||||
Db 403 IleLysThr 405

RESULT 12
S45281
coagulation factor XIIa (EC 3.4.21.38) precursor - bovine (fragment)
N:Alternate names: Hageman factor (activated)
C:Species: Bos primigenius taurus (cattle)
C>Date: 10-Apr-1995 #sequence_revision 22-Apr-1995 #text_change 21-Jan-2000
C:Accession: S45281; A61329
R:Shibuya, Y.; Semba, U.; Okabe, H.; Kambara, T.; Yamamoto, T.
Biochim. Biophys. Acta 1206, 63-70, 1994
A:Title: Primary structure of bovine Hageman factor (blood coagulation factor XII):
A:Reference number: S45281; MUID:94242782; PMID:8186251
A:Accession: S45281
A:Molecule type: mRNA
A:Residues: 1-593 <SHI>
A:Cross-references: GB:S70164
A:Note: the authors translated the codon GAG for residue 23 as Val, GAG for residue
s Pro, CTC for residue 203 as Phe, GTG for residue 247 as Leu, CCG for residue 286 a
is, and ATC for residue 505 as Leu
R:Fujikawa, K.; Walsh, K.A.; Davie, E.W.
Biochemistry 16, 2270-2278, 1977
A:Title: Isolation and characterization of bovine factor XII (Hageman factor).
A:Reference number: A61329; MUID:77182112; PMID:861210
A:Accession: A61329
A:Molecule type: protein
A:Residues: 10-16,'X',18-19;525-550 <PUJ>
C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homol
C:Keywords: blood coagulation; fibrinolysis; glycoprotein; hydrolase; monomer; plas
F:37-78/Domain: fibronectin type II repeat homology <IP2>
F:88-120/Domain: EGF homology <EGF>
F:125-160/Domain: fibronectin type I repeat homology <FB1>
F:207-287/Domain: kringle homology <KRK>
F:350-387/Domain: trypsin homology <TRY>
F:541/Active site: Ser #status predicted

Alignment Scores:
Pred. No.: 1,1e-31 Length: 593
Score: 433.00 Matches: 101
Percent Similarity: 52.14% Conservative: 45
Best Local Similarity: 36.07% Mismatches: 94
Query Match: 25.96% Indels: 40
DB: 2 Gaps: 12

US-09-735-713A-1 (1-921) x S45281 (1-593)
```

```
QY 109 TGTGGCAGAGCTCGTTAAGTACAGCTTGGAAATTTTAAACATTTTCAGTCGCATT 168
|||||
Db 336 CysGlyGlnArgLeuArgLys-----TrpLeuSerSerLeuAsnArgVal 350

QY 169 CTTGAGGAGCCAAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTGAAACAAAGG 228
|||||
Db 351 ValGlyGlyLeuValAlaLeuProGlyAlaHisProTyrIleAlaLeuTyrTrpAsp 370

QY 229 CAGAAGCATATTTGTGGAGGAAGCATGCTCTACACAGATGGGTGATCAGCGCGCTCAC 288
|||||
Db 371 Gln---HisPheCysAlaGlySerLeuIleAlaProCysTrpValLeuThrAlaAlaHis 389

QY 289 TGCATTGCAACAAACATTTGTCTACTTTGAATGTACTGTCTGGAGAGTATGACTTA 348
|||||
Db 390 CysLeuGlnAsnArgProAlaProLysGluLeuThrValValLeuGlyGlnAspArgHis 409

QY 349 AGCCAGACAGACCAGGAGAGCAAACTCTCACTATTGAAACTGTCAATCATCATCATCAT 408
|||||
Db 410 AsnGlnSerCysGluGlnCysGlnThrLeuAlaValArgAspTyrArgLeuHisGluAla 429
```



QY 808 GGTGAGGCTGGAGAAACAATGTGAGAAAAGTGTATCAAGGATCCCTGGGATCTTCA 867  
Db 996 AlaLeuProAsnArg-----ProGlyValThrAla 1005  
QY 868 GACATTAGTAAGTCTCTTCTGGATCCACGAACATC 906  
Db 1006 ArgValSerArgThrGluTrpIleGlnSerPheLeu 1018  
RESULT 14  
JC77731  
membrane-bound arginine-specific serine proteinase precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 14-Dec-2001  
R:Accession: JC7731  
R:Kishi, K.; Yamazaki, K.; Yasuda, I.; Yahagi, N.; Ichinose, M.; Tsuchiya, Y.; Athauda, J.; Biochem. 130, 425-430, 2001  
A:Title: Characterization of a membrane-bound arginine-specific serine proteinase from rat  
A:Reference number: JC7731; MUID:21421307; PMID:11530019  
A:Accession: JC7731  
A:Molecule type: mRNA  
A:Residues: 1-855 <KIS>  
A:Cross-references: DDBJ:AB049189  
A:Experimental source: strain Male, 7-week-old  
C:Comment: This enzyme, localized mainly on brushborder membranes of the intestine, part  
C:Keywords: protein digestion  
Alignment Scores:  
Pred. No.: 1.48e-30 Length: 855  
Percent Similarity: 52.67% Matches: 96  
Best Local Similarity: 36.64% Conservative: 42  
Query Match: 25.24% Mismatches: 94  
Indels: 30  
Gaps: 8  
DB:  
US-09-735-713a-1 (1-921) x JC7731 (1-855)  
QY 160 AGTCGATCTTGGAGAGCAAGTGGAGAGGTTCTATCCCTGGCAGGTATCTCTG 219  
Db 613 AlaArgValValGlyThrAsnAlaAspGluGlyGluTrpProTrpGlnValSerLeu 632  
QY 220 AAACAA----AGCAGAAGCATATTTGTGGAGAGCATCGTCTACACAGGTGGGTGATC 276  
Db 633 HisAlaLeuGlyGlnGlyHisLeuGlyHisLeuSerProAspTrpLeuVal 652  
QY 277 AGCGGGCTCACTGATTCGCAACAGAACATTTGTCTACTTTCAATCTTACT----- 330  
Db 653 SerAlaAlaHisCysPheGlnAspGluThrIlePheLeuSerProAspHisThrMetTrp 672  
QY 331 -----GCTGGAGAGTATGACTTAAGC---CAGACAGACCCAGAGAGCAAACTCTC 378  
Db 673 ThrAlaPheLeuGlyLeuAspGlnSerLysArgSerAlaSerGlyValGlnGluHis 692  
QY 379 ACTATTGAACGTGCATCATCATCCATTTCTCCACAGAAACCAATGGACTATGAT 438  
Db 693 LysLeuLysArgIleThrHisProSerPheAsn---AspPheThrPheAspTrpAsp 711  
QY 439 ATTGGCCCTTTGAAGATGGCTGGAGCCCTTCCAATTTGGCCACTTTGTGGGGCCCATATGT 498  
Db 712 IleAlaLeuLeuGluGluLysProAlaGluTrpSerThrValValArgProIleCys 731  
QY 499 CTTCCAGAGAGTGGCGGACAAATTTGAGGCTGGTTTATTTGTAACATGCGAGCTGGGCG 558  
Db 732 LeuProAspAsnThrHisValPheProAlaGlyLysAlaIleTrpValThrGlyTrpGly 751  
QY 559 CGCTTAAGTGAAGTGGGTCTCTCACAGTCTTTGCGAGGAGTGAATCTGCCTATTTTG 618  
Db 752 HisThrLysGluGlyGlyThrGlyAlaLeuLeuGlnLysGlyGluIleArgValIle 771  
QY 619 -----ACCTGGGAGAGGTGTGGCAGCTCTGTTTAACACTAAAGAGGCCCATCAGT 669  
Db 772 AsnGlnThrThrCysGluGluLeuLeuProGlnGlnIleThrProArg----- 787  
QY 670 GGAAGACCTTTCTTTGACACAGGTTTCTCTGATGGAGGAGAGAGCGCATGTCTCAGGAGAT 729

Db 788 -----MetMetCysValGlyPheLeuSerGlyGlyValAspSerCysGlnGlyAsp 804  
QY 730 TCAGGAGGTTTCACTCATGTGCGGGAATAAGAAAGGGCTGGACTCTGGCTGGTCTGACT 789  
Db 805 SerGlyGlyProLeuSerSerValGluLysAspGlyArgIlePheGlnAlaGlyValVal 824  
QY 790 TCCTGGGCTTTGGGCTGTGGTGGAGCTGGAGAAACAATGTGAGAAAGTCAATCAAGGA 849  
Db 825 SerTrpGlyGluGlyCys-----AlaGlnArgAsn 834  
QY 850 TCCCTGGGATCTTCACAGACATTTAGTAAAGTGCCTTCTGGATCCACGAACATCCAA 909  
Db 835 LysProGlyValThrArgIleProGluValArgAspTrpIleLysGlu-----Gln 852  
QY 910 ACTGGT 915  
Db 853 ThrGly 854  
RESULT 15  
JC7775  
membrane type-serine proteinase 1 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
C:Accession: JC7775  
R:Satomi, S.; Yamasaki, Y.; Tsuzuki, S.; Hitomi, Y.; Iwanaga, T.; Fushiki, T.  
Biochem. Biophys. Res. Commun. 287, 995-1002, 2001  
A:Title: A role for membrane-type serine protease (MT-SPL) in intestinal epithelial  
A:Reference number: JC7775; PMID:11573963  
A:Contents: Small intestine  
A:Accession: JC7775  
A:Molecule type: mRNA  
A:Residues: 1-855 <SAT>  
A:Cross-references: DDBJ:AB037898  
C:Comment: This enzyme, an epithelial-derived, type II integral membrane serine prot  
ciated with epithelial migration and/or cell loss not only as an upstream activator  
proteins.  
A:Gene: mt-spl  
A:Map position: basolateral cell surface  
Alignment Scores:  
Pred. No.: 1.48e-30 Length: 855  
Score: 421.00 Matches: 96  
Percent Similarity: 52.67% Conservative: 42  
Best Local Similarity: 36.64% Mismatches: 94  
Query Match: 25.24% Indels: 30  
Gaps: 8  
DB:  
US-09-735-713a-1 (1-921) x JC7775 (1-855)  
QY 160 AGTCGATCTTGGAGAGCAAGTGGAGAGGTTCTATCCCTGGCAGGTATCTCTG 219  
Db 613 AlaArgValValGlyThrAsnAlaAspGluGlyGluTrpProTrpGlnValSerLeu 632  
QY 220 AAACAA----AGCAGAAGCATATTTGTGGAGAGCATCGTCTACACAGGTGGGTGATC 276  
Db 633 HisAlaLeuGlyGlnGlyHisLeuGlyHisLeuSerProAspTrpLeuVal 652  
QY 277 AGCGGGCTCACTGATTCGCAACAGAACATTTGTCTACTTTGAATTTACT----- 330  
Db 653 SerAlaAlaHisCysPheGlnAspGluThrIlePheLeuSerProAspHisThrMetTrp 672  
QY 331 -----GCTGGAGAGTATGACTTAAGC---CAGACAGACCCAGAGAGCAAACTCTC 378  
Db 673 ThrAlaPheLeuGlyLeuAspGlnSerLysArgSerAlaSerGlyValGlnGluHis 692  
QY 379 ACTATTGAACGTGCATCATCATCCATTTCTCCACAGAAACCAATGGACTATGAT 438  
Db 693 LysLeuLysArgIleThrHisProSerPheAsn---AspPheThrPheAspTrpAsp 711  
QY 439 ATTGGCCCTTTGAAGATGGCTGGAGCCCTTCCAATTTGGCCACTTTGTGGGGCCCATATGT 498  
Db 712 IleAlaLeuLeuGluGluLysProAlaGluTrpSerThrValValArgProIleCys 731  
QY 499 CTTCCAGAGAGTGGCGGACAAATTTGAGGCTGGTTTATTTGTAACATGCGAGCTGGGCG 558  
Db 732 LeuProAspAsnThrHisValPheProAlaGlyLysAlaIleTrpValThrGlyTrpGly 751  
QY 559 CGCTTAAGTGAAGTGGGTCTCTCACAGTCTTTGCGAGGAGTGAATCTGCCTATTTTG 618  
Db 752 HisThrLysGluGlyGlyThrGlyAlaLeuLeuGlnLysGlyGluIleArgValIle 771  
QY 619 -----ACCTGGGAGAGGTGTGGCAGCTCTGTTTAACACTAAAGAGGCCCATCAGT 669  
Db 772 AsnGlnThrThrCysGluGluLeuLeuProGlnGlnIleThrProArg----- 787  
QY 670 GGAAGACCTTTCTTTGACACAGGTTTCTCTGATGGAGGAGAGAGCGCATGTCTCAGGAGAT 729

Search completed: March 25, 2003, 07:00:23  
Job time : 49.5 secs

GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model  
Run on: March 25, 2003, 05:59:35 ; Search time 182 seconds  
(without alignments)  
1348.613 Million cell updates/sec

Title: US-09-735-713A-1  
Perfect score: 1668  
Sequence: 1 atgagctctcaaatgtcttat.....acatccaaactggttaactaa 921

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+n2p model -DEV=xlh  
-Q/cn2\_1/USPTO.spool/US09735713/runat\_18032003\_124353\_23109/app\_query.fasta\_1.1095  
-DB=A\_Geneseq\_101002 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09735713 @CN1\_1.53 @runat\_18032003\_124353\_23109 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A\_Geneseq\_101002.\*  
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1626	97.5	306	22	AAE04733	Human protease hom
2	1626	97.5	306	23	AAU04760	Human protease PRT
3	1608	96.4	302	22	AAE04734	Human protease hom
4	1356	81.3	556	23	AAU02748	Amino acid sequenc
5	944	56.6	182	22	AAU02475	Toad EST encoded p
6	898	53.8	164	22	AAE04735	Human protease hom
7	567.5	34.0	913	22	ABG24246	Novel human diagno
8	567.5	34.0	1576	22	ABG07870	Novel human diagno
9	567.5	34.0	1576	22	ABG10218	Novel human diagno
10	567.5	34.0	1576	22	ABG14588	Novel human diagno
11	567.5	34.0	1576	22	ABG19887	Novel human diagno
12	561.5	33.7	1031	23	ABP60993	Novel human protei
13	489.5	29.3	802	20	AAU41710	Human PRO618 prote
14	489.5	29.3	802	21	AAU44266	Human PRO618 (UNO3
15	489.5	29.3	802	21	AAU24052	Human PRO618 prote
16	489.5	29.3	802	23	AAU02755	Amino acid sequenc
17	486	29.1	787	22	ABU71302	Drosophila melanog
18	485.5	29.1	658	22	AAE06934	Human membrane-typ
19	485.5	29.1	802	22	AAE06933	Human membrane-typ
20	476.5	28.6	235	22	AAE06932	Human membrane-typ
21	467	28.0	452	20	AAU41694	Human PRO382 prote
22	467	28.0	453	22	AAU29055	Human PRO polypt
23	467	28.0	453	22	AAE06935	Human membrane-typ
24	467	28.0	453	23	AAE23020	Human trypsin fami
25	467	28.0	454	21	AAU32246	Tumour associated
26	467	28.0	454	22	AAU68911	Human TAGD-12, pro
27	466	27.9	453	21	AAU44250	Human PRO382 (UNO3
28	456.5	27.4	327	21	AAU72093	Human serine prote
29	456.5	27.4	327	23	AAE17921	Human gene 3 encod
30	456.5	27.4	394	23	ABP41994	Human ovarian anti
31	456.5	27.4	454	23	AAE23024	Human trypsin fami
32	456.5	27.4	454	23	AAU02745	Amino acid sequenc
33	454.5	27.2	248	21	AAU43572	Human cancer assoc
34	450.5	27.0	1128	23	AAU08890	Human protease PRT
35	450.5	27.0	1128	23	AAU02739	Amino acid sequenc
36	446	26.7	243	23	AAU00516	Epithelin-like ser
37	446	26.7	309	23	AAU00531	Epithelin-like ser
38	446	26.7	818	23	AAU08753	Amino acid sequenc
39	444	26.6	414	21	AAU08912	Human secreted pro
40	444	26.6	480	21	AAU08950	Human secreted pro
41	442.5	26.5	798	15	AAU57283	Bovine enterokinas
42	441	26.4	249	23	AAE21441	Human trypsin doma
43	441	26.4	343	23	AAU078547	Human prostatas pr
44	441	26.4	343	23	ABU07285	Amino acid sequenc
45	440	26.4	416	20	AAU43325	Mouse hepsin prote

ALIGNMENTS

RESULT 1  
AAE04733  
ID AAE04733 standard; Protein; 306 AA.  
XX  
AC AAE04733;  
XX  
XX 10-SEP-2001 (first entry)  
XX  
XX Human protease homologue #1.  
DE  
XX  
KW Human; protease homologue; novel human protein; NHP; therapy;  
KW pharmacogenomic application; physiological disorder.  
XX  
OS Homo sapiens.  
XX  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 23  
FT /note "Encoded by CRA"  
FT Misc-difference 28  
FT

FT /note= "Encoded by RCT"  
 XX WO200146407-A1.  
 PN 28-JUN-2001.  
 PD  
 XX 12-DEC-2000; 2000WO-US33738.  
 PF 23-DEC-1999; 99US-0171566.  
 PR (LEXI-) LEXICON GENETICS INC.  
 XX  
 XX Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AF;  
 PI WPI; 2001-408641/43.  
 XX N-PSDB; AAD09328.  
 DR  
 XX Polynucleotide encoding novel human protease homologs, useful for  
 PT identifying agonist, antagonist or modifiers or for producing  
 PT antibodies useful in therapeutic, diagnostic and pharmacogenomic  
 PT applications -  
 XX  
 XX Claim 2; Page 28-29; 32pp; English.  
 PS  
 XX The present sequence is novel human protein (NHP),  
 CC known as human protease homologue. NHP shares structural similarity  
 CC with animal proteases, particularly trypsin-like protease such  
 CC as oviductin, plasminogen activator and human plasma kallikrein  
 CC precursor. NHP and its cDNA are useful in therapeutic, diagnostic and  
 CC pharmacogenomic applications. NHP sequences are useful for identifying  
 CC agonists, antagonists and modulators and also for producing antibodies  
 CC useful in diagnosis, drug screening, clinical trial monitoring and in  
 CC treatment of physiological disorders.  
 XX  
 XX Sequence 306 AA;  
 SQ  
 Alignment Scores:  
 Pred. No.: 2.54e-168 Length: 306  
 Score: 1626.00 Matches: 304  
 Percent Similarity: 99.35% Conservative: 0  
 Best Local Similarity: 99.35% Mismatches: 2  
 Query Match: 97.48% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-735-713A-1 (1-921) x AAE04733 (1-306)  
 QY 1 ATGAGTCTCAAAATGCTTATAGACAGGACAGCTGATTTTACTAGGATAGTCTTT 60  
 Db 1 MetSerLeuLysMetLeuLeuSerArgAsnLysLeuLeuLeuLeuGlyValPhe 20  
 QY 61 TTTGAACRAGGTAAATCTGCARCTTTTCGCTCCCAAGCTCCAGTTGGGCAGAGT 120  
 Db 21 PheGluArgGlyLysSerAlaLeuSerLeuProLysAlaProSerCysGlyClnSer 40  
 QY 121 CTGGTTAAGGTACAGCTTGAATTTTAAACATTTTTCAGTCGATTTCTGGAGGAAGC 180  
 Db 41 LeuValLysValGlnProTrpAsnTyrPheAsnLeuPheSerArgIleLeuGlyClySer 60  
 QY 181 CAAGTGGAGAGGGTTCTATCCCTGGCAGGTATCTCTGAAACAAAGCCAGACATATT 240  
 Db 61 GlnValGluLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80  
 QY 241 TGTGAGGAAGCATCTCTACACAGGTGGTGTATCGAGCGGCTCACTGCATTGCAAAAC 300  
 Db 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100  
 QY 301 AGAATCATTTGCTTACTTTGAATGTTTACTCTGAGAGTATGACTTAAGCCAGACAGAC 360  
 Db 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120  
 QY 361 CCAGGAGAGCAACTCTACATTTGAACACTGTCATCATCATCATCATCATCATCATCAT 420  
 Db 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140

QY 421 AAACCAATGGAGTATGATATTTGCCCTTTTGAAGATGGTGGAGCCTTCAATTTTGGCCAC 480  
 Db 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160  
 QY 481 TTTGTGGGGCCCATATGCTTTCAGAGCTGGGGAGCAATTTGAGCGCTGGTGTATTTGT 540  
 Db 161 PheValGlyProIleCysLeuProGlnLeuArgGluGlnPheGluAlaGlyPheIleCys 180  
 QY 541 ACAACTGCAGGCTGGGGCGCTTAACCTGAAGTGGCGCTCTCTCACAAGTCTTGAGGAA 600  
 Db 181 ThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnGlu 200  
 QY 601 GTGAATCTGCCTATTGACCTGGGAAGAGTGTGGCAGCTCTCTTAACACTAAAGAGG 660  
 Db 201 ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg 220  
 QY 661 CCCATCAGTGGAGACCTTTTTCACAGCTTTTCTGATGGAGGAGACGCGATGT 720  
 Db 221 ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys 240  
 QY 721 CAGGAGATTGAGGAGTTCATCTGTCGCGGATTAAGAAGGGCGCTGACTTGCT 780  
 Db 241 GlnGlyAspSerGlySerLeuMetCysArgAsnLysLysGlyValaTrpThrLeuAla 260  
 QY 781 GGTGTGACTTCTGGGGTTTGGCTGTGTCGAGCTGGAGAAACAATGTGAGGAAAAGT 840  
 Db 261 GlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnAsnValArgLysSer 280  
 QY 841 GATCAAGGATCCCTGGGATCTTACAGACATTAAGTAAAGTCTTCTGGATCCACGAA 900  
 Db 281 AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu 300  
 QY 901 CACATCCAACTGGTAAAC 918  
 Db 301 HisIleGlnThrGlyAsn 306  
 RESULT 2  
 AAU74760  
 ID AAU74760 standard; Protein: 306 AA.  
 XX  
 AC AAU74760;  
 XX  
 DT 09-APR-2002 (first entry)  
 XX  
 DE Human protease PRTS-20 protein sequence.  
 XX  
 KW Human; protease; PRTS; gastrointestinal; Crohn's disease; cancer;  
 KW cardiovascular; atherosclerosis; autoimmune disorder; dermatitis;  
 KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS;  
 KW cell proliferative disorder; developmental disorder; epilepsy;  
 KW Duchenne muscular dystrophy; epithelial disorder; neurological disorder;  
 KW reproductive disorder; endometriosis.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200198468-A2.  
 XX  
 PD 27-DEC-2001.  
 XX  
 PF 13-JUN-2001; 2001WO-US19178.  
 XX  
 PR 16-JUN-2000; 2000US-212336P.  
 PR 22-JUN-2000; 2000US-213955P.  
 PR 29-JUN-2000; 2000US-215396P.  
 PR 07-JUL-2000; 2000US-216821P.  
 PR 14-JUL-2000; 2000US-218946P.  
 XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 XX  
 PI Yue H, Elliott VS, Gandhi AR, Lal P, Au-young J, Tribouley CM;  
 PI Deleageane AM, Baughn MR, Nguyen DB, Lee EA, Hafalia A, Khan FA;  
 PI Wallia NK, Yao MG, Lu DAN, Patterson C, Tang YT, Walsh RT;

PI Azimzai Y, Lu Y, Ramkumar J, Xu Y, Reddy R, Das D, Kearney L;  
PI Kallick DA;  
XX WPI: 2002-090437/12.  
DR N-PSDB; ABK12903.  
XX  
PT Twenty one human proteases (referred to as PRPS-1 to PRPS-21), useful  
PT in the diagnosis, treatment and prevention of gastrointestinal (e.g.  
PT gastritis), cardiovascular (e.g. atherosclerosis) and cell  
PT proliferative (e.g. cancer) disorders -  
XX  
XX Claim 1; Page 157-158; 177pp; English.  
XX  
CC The present invention relates to twenty one new human proteases,  
CC referred to as PRPS-1 to PRPS-21. The PRPS polynucleotides and  
CC polypeptides of the invention are useful in the diagnosis, treatment and  
CC prevention of gastrointestinal e.g. gastritis, esophageal carcinoma and  
CC Crohn's disease, cardiovascular e.g. atherosclerosis, hypertension and  
CC myocardial infarction, autoimmune/inflammatory e.g. acquired  
CC immunodeficiency syndrome (AIDS), allergies and osteoarthritis, cell  
CC proliferative e.g. cancer, developmental e.g. Duchenne and Becker  
CC muscular dystrophy, epithelial e.g. dermatitis, neurological e.g.  
CC epilepsy and Alzheimer's disease and reproductive e.g. infertility and  
CC endometriosis disorders. Numerous other examples of each disorder are  
CC given in the specification. The present protein sequence represents  
CC the human protease PRPS-20 protein of the invention.  
XX  
XX  
SQ Sequence 306 AA;

Alignment Scores:  
Pred. No.: 2,54e-168 Length: 306  
Score: 1626.00 Matches: 304  
Percent Similarity: 99.35% Conservative: 0  
Best Local Similarity: 99.35% Mismatches: 2  
Query Match: 97.48% Indels: 0  
DB: 23 Gaps: 0

US-09-735-713a-1 (1-921) x AAU74760 (1-306)

QY 1 ATGAGTCTCAAAATGCTTATAAGCAGGAACAAGCTGATTTTACTAGGAATAGTCTTT 60  
DB 1 MetSerLeuLysMetLeuLeuSerArgAsnLysLeuLeuLeuLeuLeuLeuValPhe 20  
QY 61 TTGGAACRAGGTAAATCTGCARCTCTTTTCGCTCCCAAGCTCCAGTTGTGGCAGAGT 120  
DB 21 PheGluArgGlyLysSerAlaThrLeuSerLeuProLysAlaProSerCysGlyGlnSer 40  
QY 121 CTGTTAAGGTACAGCTTGAATATTATTAACATTTTACGTCGCATCTTTGGAGGAGC 180  
DB 41 LeuValLysValGlnProTrpAsnTyrPheAsnIlePheSerArgIleLeuGlyGlySer 60  
QY 181 CAAGTGGAGAGGCTTCTATCCCTGGCAGGTATCTCTGAACAAGCAGACATATT 240  
DB 61 GlnValGluLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80  
QY 241 TGTGGAGGAGCATCGCTCTACACAGTGGGTGATCACGGGGCTCACTGCATTGCAAAAC 300  
DB 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100  
QY 301 AGAACATTTGCTTACTTTGAATGTTACTGCTCGAGAGTATGACTTAAGCCAGACAGAC 360  
DB 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120  
QY 361 CCAGGAGGAGCAACTCTCACTATTGAACCTCATCATCATCATCTTCTCCACCAAG 420  
DB 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140  
QY 421 AAACCAATGGACTATGATATTCCTTTTGAAGATGGCTGGAGCTTCCCAATTTGGCCAC 480  
DB 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160  
QY 481 TTTCTGGGGCCCATATGCTTCTCCAGAGCTGGGAGCAATTTGAGCGTGTGTTATTGTT 540  
DB 141 TTTCTGGGGCCCATATGCTTCTCCAGAGCTGGGAGCAATTTGAGCGTGTGTTATTGTT

DB 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180  
QY 541 ACAACTGCAGCTGGCGCGCTTAACCTGAAGTGGCGCTCTCACAAAGTCTTGCAGGAA 600  
DB 181 ThrThrAlaGlyTrpGlyArgLeuThrGluGlyValLeuSerGlnValLeuGlnGlu 200  
QY 601 GTGAATCTGCCTATTATTGACCTGGGAAGAGTGTGTGGCAGCTCTGTTAACTAAAGAGG 660  
DB 201 ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg 220  
QY 661 CCCATCAGTGGGAAGACCTTTCTTTGGCACAGGTTTCTGTGATGGAGGAGAGCGCATGT 720  
DB 221 ProfileSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys 240  
QY 721 CAGGAGATTCAGGAGGTTCTACTGTCGCCGAATAAGAAAGGGCGCTGACCTCTGGCT 780  
DB 241 GlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAla 260  
QY 781 GGTGTGACTTCTCGGGTGTGGGTGTGGTGTGGAGCTGGAGAAACAATGTGAGGAAAAAGT 840  
DB 261 GlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnAsnValArgLysSer 280  
QY 841 GATCAAGGATCCCCGCGGATCTTCACACACATTTAGTAAAGTCTTCTCGATCCACGAA 900  
DB 281 AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu 300  
QY 901 CACATCCAAACTGCTAAC 918  
DB 301 HisIleGlnThrGlyAsn 306  
RESULT 3  
AAE04734 ID AAE04734 standard; Protein: 302 AA.  
XX AAE04734;  
XX 10-SEP-2001 (first entry)  
XX Human protease homologue #2.  
XX Human; protease homologue; novel human protein; NHP; therapy;  
XX pharmacogenomic application; physiological disorder.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
XX Misc-difference 19 /note= "Encoded by CRA"  
XX Misc-difference 24 /note= "Encoded by RCT"  
XX WO200146407-A1.  
XX 28-JUN-2001.  
XX 12-DEC-2000; 2000WO-US33738.  
XX 23-DEC-1999; 99US-0171566.  
XX (LEXI-) LEXICON GENETICS INC.  
XX Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;  
XX WPI: 2001-408641/43.  
XX N-PSDB; AAD09329.  
XX Polynucleotide encoding novel human protease homologs, useful for  
XX identifying agonist, antagonist or modifiers or for producing  
XX antibodies useful in therapeutic, diagnostic and pharmacogenomic  
XX applications -  
XX Disclosure; Page 29-30; 32pp; English.  
XX



SQ Sequence 556 AA:

Alignment Scores:  
Pred. No.: 9,24e-139 Length: 556  
Score: 1356.00 Matches: 262  
Percent Similarity: 87.95% Conservative: 8  
Best Local Similarity: 85.34% Mismatches: 29  
Query Match: 81.29% Indels: 8  
DB: 23 Gaps: 2

US-09-735-713a-1 (1-921) x AA082748 (1-556)

QY 1 ATGAGTCTCAAAATGCTTATAAGCAGAACAAAGCTGATTCTTACTAGGAATAGTCTTT 60  
Db 1 MetSerLeuLysMetLeuIleSerArgAsnLysLeuIleLeuLeuGlyIleValPhe 20  
QY 61 TTTGAACRAGGTAAATCTGCARCTCTTCGCTCCCAAGCTCCCAAGTGTGTGGCGAGAGT 120  
Db 21 PheGluArgGlyLysSerAlaThrLeuSerLeuProLysAlaProSerCysGlyGlnSer 40  
QY 121 CTGGTTAAGTACAGCTTGGATTATTTAAACATTTTCAGTCGCATCTTGGAGGAAGC 180  
Db 41 LeuValLysValGlnProIrpAsnIrpPheAsnIlePheSerArgIleLeuGlyGlySer 60  
QY 181 CAAAGTGAAGAGGTTCCCTATCCCTGCAGGTATCTCTGAACAAAGGAGGAGCATATT 240  
Db 61 GlnValGluLysGlySerTyrProIrpGlnValSerLeuLysGlnArgGlnLysHisIle 80  
QY 241 TGTGGAGGAAGCATCGTCTCACACAGTGGGTGATCAGCGCGCTCACTGCATTGCAAC 300  
Db 81 CysGlyGlySerIleValSerProGlnIrpValIleThrAlaAlaHisCysIleAlaAsn 100  
QY 301 AGAACATCTGCTACTTGTATGTTACTCTGGAGAGTATGACTTAAGCCAGACAGAC 360  
Db 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120  
QY 361 CCAGGAGAGCAAACTCTCACTATTAAGAACTGTCTATCATCATATCCACATTTCTCCACCAAG 420  
Db 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140  
QY 421 AAACCAATGACTATGATATGCGCTTTTGAAGATGGCTGGAGCCCTTCCAAATTTGGCCAC 480  
Db 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160  
QY 481 TTTGTGGGCCCATATGCTTCCAGAGCTCCGGAGCAATTTTCAGGCTGGTTTATTTCT 540  
Db 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180  
QY 541 ACAACTGCAGGCTGGGCGCGCTTAAGTGAAGTGGCGTCTCTCACAAGTCTTTCGAGGAA 600  
Db 181 ThrThrAlaGlyTrpGlyArgLeuThrGluGlyValLeuSerGlnValLeuGlnGlu 200  
QY 601 GTCAATCTGCTATTTTGACCTGGGAGAGTGTGTGGCAGCTCTGTAACTAACACTAAAGAGG 660  
Db 201 ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg 220  
QY 661 CCCATCAGTGGGAAGACCTTCTTTCACAGGTTTCTTCATCGAGGAGAGACGCATGT 720  
Db 221 ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyArgAspAlaCys 240  
QY 721 CAGGAGATTCAGAGGTTCACTCATGTGCGGAATAAGAAAGGCGCTGGACTCTGGCT 780  
Db 241 GlnGlyAspSerGlySerLeuMetCysArgAsnLysLysGlyAlaIrpAsp---Ser 259  
QY 781 GGTGTGACTTCCGCGGTTTGGGCTGTGTGAGCGCTGGAGAAACAATGTGAGGAAAGT 840  
Db 260 GlyTrpSerIleTrpGluAlaGlnValGlySerLeuGluSerArgSerArgPro 279  
QY 841 GATCAAGGATCCCT-----GGGATCTTCACAGACATTAGTAA 879  
Db 280 SerLeuGlyAsnLysValArgLeuCysLeuThrAsnAsnPhePheLysLysLeuAlaGly 299  
QY 880 GTGCTTTTCTGATCCACGAA 900

Db 300 CysGlyThrTrpCysSerGlu 306

RESULT 5  
AA024275  
ID AA024275 standard; Protein; 182 AA.  
XX  
AC AA024275;  
XX  
DT 12-OCT-2001 (first entry)  
XX  
DE Toad EST encoded protein SEQ ID NO: 1800.

XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
KW diagnostics; forensic test; gene mapping; genetic disorder;  
KW biodiversity; gene therapy; nutrition.  
XX Xenopus laevis.

XX WO200154477-A2.  
XX 02-AUG-2001.  
XX 25-JAN-2001; 2001WO-US02687.  
XX 25-JAN-2000; 2000US-0491404.  
PR 17-JUL-2000; 2000US-0617746.  
PR 03-AUG-2000; 2000US-0631451.  
PR 15-SEP-2000; 2000US-0663870.  
XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
PI Cao Y, Drmanac RA, Zhang J, Werhman T;  
XX WPI; 2001-476164/51.  
DR N-PSDB; AAH98934.

XX Isolated polypeptide for treatment of diseases, diagnostics, raising  
PT antibodies and research use -  
XX Claim 20; Page 1178-1179; 1275pp; English.

XX The present invention provides the protein and coding sequences of novel  
CC proteins from a variety of organisms, including human, dog, cat, horse,  
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)  
CC from the organism of interest. They can be used in diagnostics,  
CC forensics, gene mapping, identification of mutations, to assess  
CC biodiversity and for nutritional purposes. The present sequence is a  
CC protein of the invention.

XX SQ Sequence 182 AA;

Alignment Scores:  
Pred. No.: 5,22e-94 Length: 182  
Score: 944.00 Matches: 180  
Percent Similarity: 98.90% Conservative: 0  
Best Local Similarity: 98.90% Mismatches: 2  
Query Match: 56.59% Indels: 0  
DB: 22 Gaps: 0

US-09-735-713a-1 (1-921) x AA024275 (1-182)

QY 1 ATGAGTCTCAAAATGCTTATAAGCAGAACAAAGCTGATTCTTACTAGGAATAGTCTTT 60  
Db 1 MetSerLeuLysMetLeuIleSerArgAsnLysLeuIleLeuLeuGlyIleValPhe 20  
QY 61 TTTGAACRAGGTAAATCTGCARCTCTTCGCTCCCAAGCTCCCAAGTGTGTGGCGAGAGT 120  
Db 21 PheGluArgGlyLysSerAlaThrLeuSerLeuProLysAlaProSerCysGlyGlnSer 40

QY 121 CTGGTTAGGTACACCTTGGATTATTTTAACTATTTTCACTGCGATTCTTGGAGGAAGC 180  
Db 41 LeuVallysalGlnProTrpAsnTrpPheAsnIlePheSerArgIleLeuGlyGlySer 60  
QY 181 CAAAGTGAAGAGGTTCCATCCCTGGCAGGTATCTCTGAACAAAGGCAGACATATT 240  
Db 61 GlnValGluLysGlySerTrpProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80  
QY 241 TGTGAGGAAGCATCGTCTCACACAGTGGGTGATCACGGCGGCTCACTGCGATTGCCAAAC 300  
Db 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100  
QY 301 AGAACATGTCTACTTGAATGTTACTGCTGGAGAGTATGACTTAAGCCAGACAGAC 360  
Db 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyAspLeuSerGlnThrAsp 120  
QY 361 CCAGGAGAGCAACTCTCACTATTCAAACTCTCATATACATACACATTTCTCCACCAC 420  
Db 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140  
QY 421 AAACCAATGGACTATGATATTGCCCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGCCAC 480  
Db 141 LysProMetAspTyAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160  
QY 481 TTTGTGGGGCCCATATGCTTCCAGAGCTGCGGGAGCAATTTGAGGCTGTTTATTGT 540  
Db 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180  
QY 541 ACAACT 546  
Db 181 ThrThr 182  
RESULT 6  
AAE04735  
ID AAE04735 standard; Protein; 164 AA.  
AC AAE04735;  
XX  
XX  
DT 10-SEP-2001 (first entry)  
DE Human protease homologue #3.  
XX  
KW Human; protease homologue; novel human protein; NHP; therapy;  
KW pharmacogenomic application; physiological disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200146407-A1.  
XX  
PD 28-JUN-2001.  
XX  
PF 12-DEC-2000; 2000WO-US33738.  
XX  
PR 23-DEC-1999; 99US-0171566.  
XX  
PA (LEXI-) LEXICON GENETICS INC.  
XX  
PI Walke DW, Turner CA, Abuin A, Friedrich G, Zambrowicz B, Sands AT;  
XX  
XX WPI; 2001-408641/43.  
DR N-ESDB; AAD09330.  
XX  
XX Polynucleotide encoding novel human protease homologs, useful for  
PT identifying agonist, antagonist or modifiers or for producing  
PT antibodies useful in therapeutic, diagnostic and pharmacogenomic  
PT applications -  
XX  
PS Disclosure; Page 30-31; 32pp; English.  
XX  
XX The present sequence is novel human protein (NHP),  
CC known as human protease homologue. NHP shares structural similarity  
CC with animal proteases, particularly trypsin-like protease such  
CC as oviductin, plasminogen activator and human plasma kallikrein

CC precursor. NHP and its cDNA are useful in therapeutic, diagnostic and  
CC pharmacogenomic applications. NHP sequences are useful for identifying  
CC agonists, antagonists and modulators and also for producing antibodies  
CC useful in diagnosis, drug screening, clinical trial monitoring and in  
CC treatment of physiological disorders.  
XX  
SQ Sequence 164 AA;  
Alignment Scores:  
Pred. No.: 5,22e-89 Length: 164  
Score: 898.00 Matches: 164  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 53.84% Indels: 0  
DB: Gaps: 22  
US-09-735-713A-1 (1-921) x AAE04735 (1-164)  
QY 427 ATGGACTATGATATGCCCCCTTTTGAAGATGGCTGGAGCCTTCCAATTTGGCACCTTTGTG 486  
Db 1 MetAspTyAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHisPheVal 20  
QY 487 GGGCCCATATGCTTCCAGAGCTCGGGAGCAATTTGAGGCTGGTTTATTGTACAACT 546  
Db 21 GlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThr 40  
QY 547 GCAGCTGGGGCCGCTTAACCTGAAGTGGCGTCTCTCACAAAGTCTTCAGAGAGTGAAT 606  
Db 41 AlaGlyTrpGlyArgLeuThrGluGlyValLeuSerGlnValLeuGluValAlaAsn 60  
QY 607 CTGCTATTTTGGACCTGGGAAGAGTGTGTGGCAGCTCTGTTAAACACTAAAGAGGCCCATC 666  
Db 61 LeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArgProIle 80  
QY 667 ACTGGGAAGACCTTCTTTGACACAGGTTTCTGATGGAGGAGAGCGCATGTCAGGGA 726  
Db .81 SerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCysGlnGly 100  
QY 727 GATTGAGGAGTTCACATCATGTCCGGAATAAGAAAGGGCTGGACTCTGGCTGGTGTG 786  
Db 101 AspSerGlyGlySerLeuMetCysArgAsnLysGlyAlaThrThrLeuAlaGlyVal 120  
QY 787 ACTTCCTGGGCTTTGGGCTGTGTGGCTGGAGCTGGAGAAACAATGAGGAAAGTGATCAA 846  
Db 121 ThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSerAspGln 140  
QY 847 GGATCCCTGGGATCTTCACAGACATTTAGTAAAGTCTTCTCGATCCAGCACACATC 906  
Db 141 GlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIle 160  
QY 907 CAAACTGGTAAC 918  
Db 161 GlnThrGlyAsn 164  
RESULT 7  
ABG24246  
ID ABG24246 standard; Protein; 913 AA.  
XX  
AC ABG24246;  
XX  
DT 18-FEB-2002 (first entry)  
DE Novel human diagnostic protein #24237.  
XX  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX

```
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR N-PSDB; AAS68433.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 54605; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as 'expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 913 AA;

Alignment Scores:
Pred. No.: 1,19e-52 Length: 913
Score: 567.50 Matches: 114
Percent Similarity: 59.57% Conservative: 51
Best Local Similarity: 41.16% Mismatches: 101
Query Match: 34.02% Indels: 11
DB: Gaps: 4

US-09-735-713A-1 (1-921) x ABG24246 (1-913)
QY 109 TGTGGGAGAGTCTGGTTAAGGTA-----CAGCCTTGGATATTTTAACTTTTC 159
Db 9 CysGlyIleArgMetValAsnMetLysSerLysGluProAlaValGlySerArgPhePhe 28
QY 160 AGTCGCATTCTTGGAGACCCCAAGTGGAGAGGTTCTCTATCCCTGGCAGGATCTCTG 219
Db 29 SerArgIleSerTrpArgAsnSerThrValThrGlyHisProTrpGlnValSerLeu 48
QY 220 AAACAAGGCGAGCAATTTTGGAGGAGCATCTCTCCACACAGTCAGTGGGTGATCACG 279
Db 49 LysSerAspGluHisPheCysGlyGlySerLeuIleGlnAspArgValValThr 68
QY 280 GCGGCTCAGTCATT-----GCAACAGAAACATTGTCTACTTTGAATGTACGCT 333
Db 69 AlaAlaHisCysLeuAspSerLeuSerGluLysGlnLeuLysAsnIleThrValThrSer 88
QY 334 GGAGAGTATCACTTAAGCCAGACAGACCAGGAGGAGCAAACTCTCACTATTGAACGTGC 393
Db 89 GlyGluTrpSerLeuPheGlnGlyAspLysGlnGlnAsnIleProValSerLysIle 108

394 ATCATACATCCACATTTCTCCACCAAGAAACCAATGAGCTATGATATTTGCCCTTTTGAAG 453
109 IleThrHisProGluTrpAsnSerArgGluTrpMetSerProAspIleAlaLeuLeuTyr 128
454 ATGGCTGGAGCCTTCCAATTTGGCCACTTTTGGGGCCCATATCTCTTCCAGAGTCGCGG 513
129 LeuLysHisLysValLysPheGlyAsnAlaValGlnProIleCysLeuProAspSerAsp 148
514 GAGCAATTTGAGGCTGGTTTATTTGTACAACTGCAGGCTGGGGCCCTTAACTGAAGGT 573
149 AspLysValGluProGlyIleLeuCysLeuSerSerGlyTrpGlyLysIleSerLysThr 168
574 GCGCTCTCTCAACAAGTCTTTCAGGAGTGAATCTGCCTATTTTGTACCTGGGAGAGTGT 633
169 SerGluTrpSerAsnValLeuGlnGluMetGluLeuProIleMetAspArgAlaCys 188
634 GTGGCAGCTCTGTTAACACTAAAGAGGCCCATCAGTGGGAGACCTTCTTTTCACAGGT 693
189 AsnThrValLeuLysSerMetAsnLeuProProLeuGlyArgThrMetLeuCysAlaGly 208
694 TTTCCTGATGGAGGAGAGAGCATGTCAGGGAGATTCAGGAGTTCACCTCATGTGCGG 753
209 PheProAspTrpGlyMetAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysArg 228
754 AATAAGAAAGGGCCTGGACTCTGGCTGGTGTGACTTCTCTCTGGGTTTGGGCTGTGGTCA 813
229 ArgGlyGlyGlyIleTrpIleLeuAlaGlyIleThrSerTrpValAlaGlyCysAlaGly 248
814 GGC-----TGGAGAAACAATGTGAGGAAAGATGATCAAGGATCCCTCGGATCTTC 864
249 GlySerValProValArgAsnAsnHisValLysAla-----SerLeuGlyIlePhe 265
865 ACAGACATTAGTAAGTGTCTTCTCGATCCAGCAACACATCCAAACTGCT 915
266 SerLysValSerGluLeuMetAspPheIleThrGlnAsnLeuPheThrGly 282

RESULT 8
ABG07870
ID ABG07870 standard; Protein; 1576 AA.
XX
AC ABG07870;
XX
DT 13-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #7861.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR N-PSDB; AAS72057.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID No 38229; 103pp; English.
```

XX The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 1576 AA;

Alignment Scores:  
 Pred. No.: 1.48e-52 Length: 1576  
 Score: 567.50 Matches: 114  
 Percent Similarity: 59.57% Conservative: 51  
 Best Local Similarity: 41.16% Mismatches: 101  
 Query Match: 34.02% Indels: 11  
 DB: 22 Gaps: 4

US-09-735-713A-1 (1-921) x ABG07870 (1-1576)

QY 109 TGTGGCAGAGCTCGTTAAGTA-----CAGCCTTGGAAATATTAAACATTTTC 159  
 DB 848 CysGlyIleArgMetValAsnMetLysSerLysGluProAlaValGlySerArgPhePhe 867  
 QY 160 AGTCCGCTCTGGAGGAGCAAGTGGAGAGGTTCTCTATCCCTGGCAGGTATCTCTG 219  
 DB 868 SerArgIleSerSerTrpArgAsnSerThrValThrGlyHisProTrpGlnValSerLeu 887  
 QY 220 AAACAAGGAGAGCATATTGTGGAGGAGCATCTCACCACAGTGGGTGATCAGC 279  
 DB 888 LysSerAspGluHisHisPheCysGlySerLeuIleGlnGluAspArgValThr 907  
 QY 280 GCGGCTCACTGATT-----GCAACAGAACATTTGCTCTACTTTGAATTTACTGCT 333  
 DB 908 AlaAlaHisCysLeuAspSerLeuSerGluLysGlnLeuLysAsnIleThrValThrSer 927  
 QY 334 GGAGAGTATGACTTAAGCCAGACACACCCAGGAGAGCAAACTCTCAGTATTGAAACTGTC 393  
 DB 928 GlyGluTrpSerLeuPheGlnLysAspLysGlnGluGlnAsnIleProValSerLysIle 947  
 QY 394 ATCATACATCCACATTTCTCCACCAAGAACCAATGAGTATGATTTGCCCTTTTGAAG 453  
 DB 948 IleThrHisProGluTrpAsnSerArgGluTrpMetSerProAspIleAlaLeuLeuTrp 967  
 QY 454 ATGGCTGGAGCCTTCCATTTGGCCACTTTTGGGGGCCATATGCTTCCAGAGCTGGCG 513  
 DB 968 LeuLysHisLysValLysPheGlyAsnAlaValGlnProIleCysLeuProAspSerAsp 987  
 QY 514 GAGCAATTTAGGCTGGTTTATTGTACAACTGCGAGCTGGGGCGCTTAACTGAAGGT 573  
 DB 988 AspLysValGluProGlyIleLeuCysLeuSerSerGlyTrpGlyLysIleSerLysThr 1007  
 QY 574 GCGGCTCTCACAGTCTTCAGGAAGTGAATTCCTATTTTGAACCTGGGAAGAGTGT 633  
 DB 1008 SerGluTrpSerAsnValLeuGlnGluMetGluLeuProIleMetAspArgAlaCys 1027  
 QY 634 GTGGCAGCTCTGTAAACATAAGAGGCCCATCAGTGGGAAGACCTTTCTTTCACAGGT 693

Db 1028 AsnThrValLeuLysSerMetAsnLeuProProLeuGlyArgThrMetLeuCysAlaGly 1047  
 QY 694 TTTCTGTATGGAGGAGAGCATGTCTCAGGAGATTTCAGGAGTTCATCTATGTCGGG 753  
 Db 1048 PheProAspTrpGlyMetAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysArg 1067  
 QY 754 AATAAGAAAGGGCCTGGACTTGGCTGCTGTGACTTCTCTGGGGTTTGGCTGGTCTGA 813  
 Db 1068 ArgGlyGlyIleTrpIleLeuAlaGlyIleThrSerTrpValAlaGlyCysAlaGly 1087  
 QY 814 GGC-----TGGAGAAACAATGTGAGGAAACTGATCAAGGATCCCTGGGATCTTC 864  
 Db 1088 GlySerValProValArgAsnHisValLysAla-----SerLeuGlyIlePhe 1104  
 QY 865 ACAGACATTAGTAAGTCTTCTTCGTGGATCCAGCAACACATCCAACTGGT 915  
 Db 1105 SerLysValSerGluLeuMetAspPheIleThrGlnAsnLeuPheThrGly 1121  
 RESULT 9  
 ABG10218  
 ID ABG10218 standard; Protein: 1576 AA.  
 XX  
 AC ABG10218;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #10209.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001WO-US08631.  
 XX  
 PR 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Drmanac RT, Liu C, Tang YT;  
 XX  
 DR WPI; 2001-639362/73.  
 DR N-PSDB; AAS74405.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 20; SEQ ID No 40577; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human

CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

```
XX
SQ Sequence 1576 AA;

Alignment Scores:
Pred. No.: 1.48e-52 Length: 1576
Score: 567.50 Matches: 114
Percent Similarity: 59.57% Conservative: 51
Best Local Similarity: 41.16% Mismatches: 101
Query Match: 34.02% Indels: 11
DB: 22 Gaps: 4

US-09-735-713A-1 (1-921) x ABG10218 (1-1576)
QY 109 TGTGGCAGAGTCTGTTAAGTA-----CAGCCTTGGAAATATTTTAACATTTTC 159
Db CysGlyIleArgMetValAsnMetLysSerLysGluProAlaValGlySerArgPhePhe 867
QY 160 AGTCGATCTCTGGAGAGCCCAAGTGGAGAGGGTTCCTCCCTGGCAGGTATCTCTG 219
Db SerArgIleSerSerTrpArgAsnSerThrValThrGlyHisProTrpGlnValSerLeu 887
QY 220 AAACAAAGCAGAGCATATTTCTGGAGGAGCATGCTCCACCACAGTGGGTGATCAG 279
Db LysSerAspGluHisPheCysGlySerLeuIleGlnAspArgValValThr 907
QY 280 GCGGCTCACTGCATT-----GCAACAGAAACATTTGTCTACTTTGAATGTACTGCT 333
Db AlaAlaHisCysLeuAspSerLeuSerGluLysGlnLeuLysAsnIleThrValThrSer 927
QY 334 GGAGATGATGACTTAGCCAGCAGACACCAGGAGCAACTCTCACTATTGAACTGTC 393
Db GlyGluTrpSerLeuPheGlnLysAspLysGlnGlnAsnIleProValSerLysIle 947
QY 394 ATCATACATCCACATTTCTCCACCAAGAAACCAATGATATATTTGCCCTTTTGAAG 453
Db IleThrHisProGluTrpAsnSerArgGluTrpMetSerProAspIleAlaLeuLeuTyr 967
QY 454 ATGGCTGGAGCCCTCAATTTGGCCACTTTGTGGGCCCATATGCTTCCAGAGCTGCGG 513
Db LeuLysHisLysValLysPheGlyAsnAlaValGlnProIleCysLeuProAspSerAsp 987
QY 514 GACCAATTTGAGCCTGGTTTATTCTACACTGCGAGGTGGGGCCGCTTAAGTGAAGT 573
Db AspLysValGluProGlyLleLeuCysLeuSerSerGlyTrpGlyLysIleSerLysThr 1007
QY 574 GCGCTCTCTCAACATCTTGCAAGAGTGAATCTGCCTATTATTGACCTGGGAAGAGTGT 633
Db SerGluTrpSerAsnValLeuGlnGluMetGluLeuProIleMetAspArgAlaCys 1027
QY 634 GTGGCAGCTCTGTAACTAAAGAGGCCCATCAGTGGGAAGACCTTTCTTTCACAGGT 693
Db AsnThrValLeuLysSerMetAsnLeuProProLeuGlyArgThrMetLeuCysAlaGly 1047
QY 694 TTTCTCTGATGGAGGAGAGCATGTCAGGAGATTTCAGGAGTTCACTCATGTCGCGG 753
Db PheProAspTrpGlyMetAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysArg 1067
QY 754 AATAAGAAAGGCCCTGGACTCTGGCTGTGTGACTCTCTCTGGGTTTGGCTGCTGCTGA 813
Db ArgGlyGlyGlyIleIleIleIleAlaGlyIleIleIleSerTrpValAlaGlyCysAlaGly 1087
QY 814 GGC-----TGGAGAAACAATGTGAGGAAAGATGATCAAGGATCCCTGGGATCTTC 864
Db GlySerValProValArgAsnAsnHisValLysAla-----SerLeuGlyIlePhe 1104
QY 865 ACAGACATAGTAAGTGTCTTCTCGATCCACGACACATCCAAACTGTT 915
Db SerLysValSerGluLeuMetAspPheIleThrGlnAsnLeuPheThrGly 1121
```

RESULT 10

ABG14588

ID ABG14588 standard; Protein; 1576 AA.

XX AC ABG14588;

XX DT 18-FEB-2002 (first entry)

XX DE Novel human diagnostic protein #14579.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX KW food supplement; medical imaging; diagnostic; genetic disorder.

XX OS Homo sapiens.

XX PN WO2001/5067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US08631.

XX PR 31-MAR-2000; 2000US-0540217.

XX PR 23-AUG-2000; 2000US-0649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR N-PSDB; AAS78775.

XX PT New isolated polynucleotide and encoded polypeptides, useful in

XX PT diagnostics, forensics, gene mapping, identification of mutations

XX PT responsible for genetic disorders or other traits and to assess

XX PT biodiversity -

XX PS Claim 20; SEQ ID No 44947; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and

XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

XX CC and gene mapping, and in recombinant production of (II). The

XX CC polynucleotides are also used in diagnostics as expressed sequence tags

XX CC for identifying expressed genes. (I) is useful in gene therapy techniques

XX CC to restore normal activity of (II) or to treat disease states involving

XX CC quantitating a polypeptide in tissue, as molecular weight markers and as

XX CC a food supplement. (II) and its binding partners are useful in medical

XX CC imaging of sites expressing (II). (I) and (II) are useful for treating

XX CC disorders involving aberrant protein expression or biological activity.

XX CC The polypeptide and polynucleotide sequences have applications in

XX CC diagnostics, forensics, gene mapping, identification of mutations

XX CC responsible for genetic disorders or other traits to assess biodiversity

XX CC and to produce other types of data and products dependent on DNA and

XX CC amino acid sequences. ABG0010-ABG30377 represent novel human

XX CC diagnostic amino acid sequences of the invention.

XX CC Note: The sequence data for this patent did not appear in the printed

XX CC specification, but was obtained in electronic format directly from WIPO

XX CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 1576 AA;

Alignment Scores:

Pred. No.: 1.48e-52 Length: 1576

Score: 567.50 Matches: 114

Percent Similarity: 59.57% Conservative: 51

Best Local Similarity: 41.16% Mismatches: 101

Query Match: 34.02% Indels: 11

DB: 22 Gaps: 4

US-09-735-713A-1 (1-921) x ABG14588 (1-1576)

QY 109 TGTGGCAGAGTCTGTTAAGTA-----CAGCCTTGGAAATATTTTAACATTTTC 159



```
QY 394 ATCATACATGCACATTCTCCACCAAGAAACCAATGGACTATGATATTGCCCTTTTGAAG 453
   ||| ||||| : : : : : : : : : ||| ||||| ||||| |||||
Db 948 lIeThrHisProGluTyAsnSerArgGluTyMetSerProAspIleAlaLeuLeuTyR 967
QY 454 ATGGCTGGAGCCTTCCAAATTTGGCCACATTTGGGGCCCATATGCTCCAGAGTCGGG 513
   : : : : : : : : : ||| ||||| ||||| ||||| : : :
Db 968 LeuLysHisLysValLysPheGlyAsnAlaValGlnProIleCysLeuProAspSerAsp 987
QY 514 GAGCAATTTGAGGCTGGTGTATTTGTACAACCTGAGGCTGGGCGCTTAACCTGAAGT 573
   : : : : : ||| ||| : : : : : : : : : : : : : : : : : :
Db 988 AspLysValGluProGlyIleLeuCysLeuSerGlyTrpGlyLysIleSerLysThr 1007
QY 574 GCGCTCTCTCACAGTCTTGCAGGAAGTCAATCTCCCTATTGACCTGGGAAGAGTGT 633
   ||| ||||| ||||| : : : : : : : : : : : : : : : : :
Db 1008 SerGluTyTrSerAsnValLeuGlnGluMetGluLeuProIleMetAspArgAlaCys 1027
QY 634 GTGGGAGCTCTGTACACTAAAGAGGCCCATCAGTGGGAAGACCTTTCTTGCACAGT 693
   : : : : : : : : : ||| ||||| ||||| ||||| |||||
Db 1028 AsnThrValLeuLysSerMetAsnLeuProProLeuGlyArgThrMetLeuCysAlaGly 1047
QY 694 TTTCTGTGATGAGGAGACACATGTCAGGAGATTCAGGAGGTTCACTCATGTGCGG 753
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1048 PheProAspTrpGlyMetAspLysGlnGlyAspSerGlyGlyProLeuValCysArg 1067
QY 754 AATAAGAAAGGGCGCTGACTGCTGGTGTGACTTCTCTGGGGTTGGGCTCTGTCGA 813
   ||| ||| ||||| ||||| : : : : : : : : : |||||
Db 1068 ArgGlyGlyIleTrpIleLeuAlaGlyIleThrSerTrpValAlaGlyCysAlaGly 1087
QY 814 GGC-----TGAGAAACAATGTGAGGAAGATGATCAAGGATCCCTGGGATCTTC 864
   ||| ||||| ||||| ||||| : : : : : |||||
Db 1088 GlySerValProValArgAsnAsnHisValLysAla-----SerLeuGlyIlePhe 1104
QY 865 ACAGACATTAGTAAGTGTCTTCTCGATCCACGACACATCCAACTGCT 915
   : : : : : : : : : : : : : : : : : : : : : : :
Db 1105 SerLysValSerGluLeuMetAspPheIleThrGlnAsnLeuPheThrGly 1121

RESULT 12
ID ABP60993
XX ABP60993 standard; Protein; 1031 AA.
AC ABP60993;
XX
XX
XX 10-SEP-2002 (first entry)
XX
XX
XX Novel human protein. SEQ ID 80.
XX
KW Human; cytostatic; vulnery; antiarteriosclerotic; antiparkinsonian;
KW nootropic; neuroprotective; immunosuppressive; haemostatic;
KW antiinflammatory; cardiant; antiulcer; virucide; antithyroid;
KW cerebroprotective; anorectic; metabolic; vaccine; cancer; infection;
KW wound healing disorders; atherosclerosis; Parkinson's disease;
KW Alzheimer's disease; autoimmune disorder; haematopoietic disorder;
KW inflammation; neoplastic disease; nervous system disorder;
KW cardiovascular disorders; pancreatitis; respiratory disorder;
KW hyperproliferation; systemic autoimmune disease; hyper-immunity;
KW developmental abnormality; gastrointestinal ulceration; neuropathy;
KW haematological disease; metabolic disease; sperm dysfunction;
KW thyroid disorder; hypothyroidism; brain damage; colitis;
KW cone photo- transduction deficiency; neurological disease; stroke;
KW anglogenesis; ovulation disorder; spinal cord; thyroid gland; heart;
KW trachea; thymus; lymph node; muscular system; obesity; anorexia;
KW growth abnormality; precocious puberty.
XX
OS Homo sapiens.
XX
XX WO200250105-A1.
XX
XX 27-JUN-2002.
XX
XX 17-DEC-2001; 2001WO-US49232.
XX
XX 19-DEC-2000; 2000US-256710P.
```

```
PR 20-DEC-2000; 2000US-257048P.
PR 09-JAN-2001; 2001US-260482P.
PR 30-JAN-2001; 2001US-264922P.
PR 06-FEB-2001; 2001US-266797P.
PR 19-MAR-2001; 2001US-276988P.
PR 04-APR-2001; 2001US-281535P.
PR 08-MAY-2001; 2001US-289622P.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
PA (SMIK ) SMITHKLINE BEECHAM PLC.
PA (GLAX ) GLAXO GROUP LTD.
XX
PI Agarwal P, Birkeland M, Cogswell JP, Kabnick KF, Lai Y;
PI Martensen SA, Rizvi SK, Smith RF, Strum JC, Xie Q;
XX
XX WPI; 2002-508784/54.
XX N-PSDB; ABQ86158.
XX
PT Secreted proteins and polynucleotides useful as vaccines for preventing
PT or treating various diseases e.g. cancer, wounds, atherosclerosis,
PT Parkinson's disease, Alzheimer's disease, infection, autoimmune
PT disorder -
XX
XX Claim 1(a); Page 299-301; 335pp; English.
XX
XX The invention relates to an isolated polypeptide with signal sequences
XX which allow it to be secreted extracellularly or membrane associated.
XX The activity of polypeptides of the invention may be described as,
XX cytostatic, vulnery, antiarteriosclerotic, antiparkinsonian, nootropic,
XX neuroprotective, immunosuppressive, haemostatic, antiinflammatory,
XX cardiant, antiulcer, virucide, antithyroid, cerebroprotective, anorectic,
XX and metabolic. Polypeptides and polynucleotides of the invention are
XX useful in the treatment, or as a vaccine in the prevention of, cancer,
XX wound healing disorders, infection, atherosclerosis, Parkinson's disease,
XX and Alzheimer's disease, autoimmune disorder, haematopoietic disorder,
XX inflammation, neoplastic diseases, nervous system related disorders and
XX cardiovascular disorders, pancreatitis, respiratory disorder,
XX hyperproliferation, systemic autoimmune disease, hyper-immunity,
XX developmental abnormality, gastrointestinal ulceration, neuropathy,
XX haematological diseases, metabolic diseases, sperm dysfunction, thyroid
XX disorders e.g. hypothyroidism, brain damages, colitis, cone photo-
XX transduction deficiency, neurological diseases, stroke, anglogenesis,
XX ovulation disorders, diseases in the spinal cord, thyroid gland, heart,
XX trachea, thymus, lymph node and muscular system, obesity, anorexia,
XX growth abnormalities, and alleviation of precocious puberty. The
XX sequences given in records ABP60965-ABP61019 represent novel human
XX proteins of the invention.
XX
SQ Sequence 1031 AA;
XX
XX Alignment Scores:
XX Pred. No.: 5,62e-52 Length: 1031
XX Score: 561.50 Matches: 113
XX Percent Similarity: 59.42% Conservative: 51
XX Best Local Similarity: 40.94% Mismatches: 101
XX Query Match: 33.66% Indels: 11
XX DB: 23 Gaps: 4
XX
XX US-09-735-713A-1 (1-921) x ABP60993 (1-1031)
QY 109 TGTGGCAGAGTCTGGTTAAGTA-----CAGCCTTGGAAATATTTTAAACATTTTC 159
   ||||| : : : : : : : : : |||||
Db 9 CysGlyIleArgMetValAsnMetLysSerLysGluProAlaValGlySerArgPhePhe 28
QY 160 AGTCGATTCTTGGAGGAGCCCAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTG 219
   ||||| ||||| : : : : : : : : : |||||
Db 29 SerArgIleSerSerTrpArgAsnSerThrValThrGlyHisProTrpGlnValSerLeu 48
QY 220 AAACAAGGCACAGCATATTTGTGAGGAGCACTGCTCTACCACAGTGGGTGATCAGC 279
   ||| : : : ||| ||||| ||||| : : : : : |||||
Db 49 LysSerAspGluHisPheCysGlyGlySerLeuIleGlnGluAspArgValThr 68
QY 280 GCGGCTCACTGCATT-----GCAACACAAACATTTGTCTACTTTGAATGTTACTGCT 333
```

```
Db 69 AlalaahiscysLeuAspSerLeuSerGluLysGlnLeuLysAsnIleThrValThrSer 88
      |||||
QY 334 GGAGCATGATGACCTTAAGCCAGACAGACCCAGGAGAGAACTCTCAGTATTAAGAACTGTC 393
      |||||
Db 89 GlyGluThrSerLeuPheGlnLysAspLysGlnGlnAsnIleProValSerLysIle 108
      |||||
QY 394 ATCATACATCCACATTTCTCCACCAAGAACCAATGACATGATATTTGCCCTTTTGAAG 453
      |||||
Db 109 IleThrHisProGluThrAsnSerArgLutyrMetSerProAspIleAlaLeuLeuThr 128
      |||||
QY 454 ATGCTCGAGCGCTTCCAAATTTGGCCACTTTTGGGGCCCATATGCTTCCAGAGCTCGG 513
      |||||
Db 129 LeuLysHisLysValLysPheGlyAsnAlaValGlnProIleCysLeuProAspSerAsp 148
      |||||
QY 514 GAGCAATTTGAGGCTGCTTTTATTTAGCACTGCACTGCGGCGGCGCTTAAGTGAAGT 573
      |||||
Db 149 AspLysValGluProGlyIleLeuCysLeuSerSerGlyTrpGlyLysIleSerLysThr 168
      |||||
QY 574 GGCCTCCTCTCACAAAGTCTTGCAGAAAGTGAATCTGCTATTTTGACCTGGGAAGAGTGT 633
      |||||
Db 169 SerGluThrSerAsnValLeuGlnMetGluLeuProIleMetAspAspArgAlaCys 188
      |||||
QY 634 GTGCGAGCTGTGTTAACTAAAGAGAGCCCATCATGCTGGGAGAGACCTTCTTTCACACAGT 693
      |||||
Db 189 AsnThrValLeuLysSerMetAsnLeuProLeuGlyArgThrMetLeuCysAlaGly 208
      |||||
QY 694 TTTCTCATGAGGAGAGAGCGCATGTCAGGAGATTCAGAGGTTTCACTCATGTCGCGG 753
      |||||
Db 209 PheProAspTrpGlyMetAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysArg 228
      |||||
QY 754 AATAAGAAAGGGCGCTGACTCTGGCTGTGCTGACTTCCCTGGGCTTTGGGCTGTGGTCGA 813
      |||||
Db 229 ArgGlyGlyGlyIleThrPheLeuAlaGlyIleThrSerTrpValAlaGlyCysAlaGly 248
      |||||
QY 814 GGC-----TGGAGAACATGTGAGGAAAGTATCAAGATCCCTCGGATCTTC 864
      |||||
Db 249 GlySerValProValArgAsnAsnHisValLysAla-----SerLeuGlyIlePhe 265
      |||||
QY 865 ACACACATTAAGTAAAGTCTTCTGCTGATCCACGACACATCCAAACT 912
      |||||
Db 266 SerLysValSerGluLeuMetAspPheIleThrGlnAsnLeuPheThr 281
      |||||
RESULT 13
AAY41710
ID AAY41710 standard; Protein; 802 AA.
XX
AC AAY41710;
XX
DT 07-DEC-1999 (first entry)
XX
DE Human PRO618 protein sequence.
XX
KW Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
KW probe; blood coagulation disorder; cancer; cellular adhesion disorder;
KW secreted protein; transmembrane protein.
XX
OS Homo sapiens.
XX
PN W09946281-A2.
XX
PD 16-SEP-1999.
XX
PF 08-MAR-1999; 99WO-US05028.
XX
PR 10-MAR-1998; 98US-0077450.
PR 11-MAR-1998; 98US-0077632.
PR 11-MAR-1998; 98US-0077641.
PR 11-MAR-1998; 98US-0077649.
PR 12-MAR-1998; 98US-0077791.
PR 13-MAR-1998; 98US-0078004.
PR 17-MAR-1998; 98US-0040220.
PR 20-MAR-1998; 98US-0078886.
PR 20-MAR-1998; 98US-0078910.
PR 20-MAR-1998; 98US-0078936.
PR 20-MAR-1998; 98US-0078939.
PR 25-MAR-1998; 98US-0079294.
PR 26-MAR-1998; 98US-0079656.
PR 27-MAR-1998; 98US-0079663.
PR 27-MAR-1998; 98US-0079664.
PR 27-MAR-1998; 98US-0079689.
PR 27-MAR-1998; 98US-0079728.
PR 27-MAR-1998; 98US-0079786.
PR 30-MAR-1998; 98US-0079920.
PR 31-MAR-1998; 98US-0079923.
PR 31-MAR-1998; 98US-0080105.
PR 31-MAR-1998; 98US-0080107.
PR 31-MAR-1998; 98US-0080165.
PR 31-MAR-1998; 98US-0080194.
PR 01-APR-1998; 98US-0080327.
PR 01-APR-1998; 98US-0080328.
PR 01-APR-1998; 98US-0080333.
PR 01-APR-1998; 98US-0080334.
PR 08-APR-1998; 98US-0081049.
PR 08-APR-1998; 98US-0081070.
PR 08-APR-1998; 98US-0081071.
PR 09-APR-1998; 98US-0081195.
PR 09-APR-1998; 98US-0081203.
PR 09-APR-1998; 98US-0081229.
PR 15-APR-1998; 98US-0081817.
PR 15-APR-1998; 98US-0081838.
PR 15-APR-1998; 98US-0081952.
PR 15-APR-1998; 98US-0081955.
PR 21-APR-1998; 98US-0082568.
PR 21-APR-1998; 98US-0082569.
PR 22-APR-1998; 98US-0082700.
PR 22-APR-1998; 98US-0082704.
PR 22-APR-1998; 98US-0082804.
PR 23-APR-1998; 98US-0082767.
PR 23-APR-1998; 98US-0082796.
PR 27-APR-1998; 98US-0083336.
PR 28-APR-1998; 98US-0083322.
PR 29-APR-1998; 98US-0083392.
PR 29-APR-1998; 98US-0083495.
PR 29-APR-1998; 98US-0083496.
PR 29-APR-1998; 98US-0083499.
PR 29-APR-1998; 98US-0083500.
PR 29-APR-1998; 98US-0083545.
PR 29-APR-1998; 98US-0083554.
PR 29-APR-1998; 98US-0083558.
PR 29-APR-1998; 98US-0083559.
PR 30-APR-1998; 98US-0083742.
PR 05-MAY-1998; 98US-0084366.
PR 06-MAY-1998; 98US-0084414.
PR 06-MAY-1998; 98US-0084441.
PR 07-MAY-1998; 98US-0084598.
PR 07-MAY-1998; 98US-0084600.
PR 07-MAY-1998; 98US-0084627.
PR 07-MAY-1998; 98US-0084637.
PR 07-MAY-1998; 98US-0084639.
PR 07-MAY-1998; 98US-0084640.
PR 07-MAY-1998; 98US-0084643.
PR 13-MAY-1998; 98US-0085323.
PR 13-MAY-1998; 98US-0085338.
PR 13-MAY-1998; 98US-0085339.
PR 15-MAY-1998; 98US-0085573.
PR 15-MAY-1998; 98US-0085579.
PR 15-MAY-1998; 98US-0085580.
PR 15-MAY-1998; 98US-0085582.
PR 15-MAY-1998; 98US-0085689.
PR 15-MAY-1998; 98US-0085697.
PR 15-MAY-1998; 98US-0085700.
PR 15-MAY-1998; 98US-0085704.
PR 18-MAY-1998; 98US-0086023.
PR 22-MAY-1998; 98US-0086392.
PR 22-MAY-1998; 98US-0086414.
```



DR N-PSDB; AAC78494.  
XX Novel PRO polypeptides and polynucleotides used in detection methods,  
PT to target bioactive molecules to specific cells, and to modulate  
PT cellular activities -  
XX  
XX  
XX Claim 12; Fig 63; 636pp; English.  
XX  
CC AAC78458 to AAC78599 represent polynucleotide and EST (expressed  
CC sequence tag) sequences which encode secreted or transmembrane PRO  
CC polypeptides. The PRO polynucleotides and polypeptides have cytostatic  
CC activity. The polynucleotides and polypeptides can be used for detecting  
CC the presence of PRO polypeptides in samples, for linking bioactive  
CC molecules to cells and for modulating biological activities of cells,  
CC using the polypeptides for specific targeting. The polypeptide targeting  
CC can be used to kill the target cells, e.g. for the treatment of cancers.  
CC The polypeptide pairs provide specific targeting of bioactive molecules  
CC to cells. AAC78600 to AAC78987 represent PCR primers and probes used in  
CC the isolation of the PRO polynucleotide sequences.  
XX  
SQ Sequence 802 AA;  
  
Alignment Scores:  
Pred. No.: 3.64e-44 Length: 802  
Score: 489.50 Matches: 99  
Percent Similarity: 56.75% Conservative: 44  
Best Local Similarity: 39.29% Mismatches: 86  
Query Match: 29.35% Indels: 23  
DB: 21 Gaps: 5  
  
US-09-735-713A-1 (1-921) x AAB44266 (1-802)  
QY 160 AGTCGATCTTGGAGGAGCAAGTGGAGAGGTTCTATCCCTGGCAGGTATCTCTG 219  
DB 566 SerArgIleValGlyAlaValSerSerGluGlyGluTrpProTrpGlnAlaSerLeu 585  
QY 220 AAACAAGGAGCAAGCATATTGTGGAGAGCATGCTCTCACCACAGTGGGTGATCAGC 279  
DB 586 GlnValArgGlyArgHisIleCysGlyGlyAlaLeuIleAlaAspArgTrpValIleThr 605  
QY 280 GCGGCTCACTGCATGCAACAGAAACATTTGTCTACTTTG-----AATGTTACTGCT 333  
DB 606 AlaAlaHisCysPheGlnGluAspSerMetAlaSerThrValLeuTrpThrValPheLeu 625  
QY 334 GGAGAGTAGTACTTAAGCAGACAGACCCAGGAGAGCAACTCTCACTATTGAAACGTGC 393  
DB 626 GlyLysValTrpGlnAsnSerArgTrpProGlyGluValSerPheLysValSerArgLeu 645  
QY 394 ATCAATACATCCACATTTCTCCACCAAGAAACCAATGGACTATGATATTCCTTTTGAAG 453  
DB 646 LeuLeuHisPro--TyrHisGluGluAspSerHisAspTyrAspValAlaLeuLeuGln 664  
QY 454 ATGGCTGGAGCCTCCCAATTTGGCCACTTTGTGGGGCCATATGCTTCCAGAGCTCGGG 513  
DB 665 LeuAspHisProValValArgSerAlaAlaValArgProValCysLeuProAlaArgSer 684  
QY 514 GAGCAATTTAGGCTGGTTTATTGTTACAACTCAGCGCTGGCGCGCTTAACCTGAAGGT 573  
DB 685 HisPheGluProGlyLeuHisCysTrpIleThrGlyTrpGlyAlaLeuArgGluGly 704  
QY 574 GCGCTCTCTCACAGTCTTTCAGGAAGTGAATCTGCTATTTTGAAGTGGGAAGAGTGT 633  
DB 705 GlyProIleSerAsnAlaLeuGlnLysValAspValGlnLeuIleProGlnAspLeuCys 724  
QY 634 GTGGCAGCTCTGTTAACTAAAGAGGCCCATCAGTGGGAAGACCTTTCTTTTGCACAGGT 693  
DB 725 SerGluAlaTrpArgTyrGlnValThrPro-----ArgMetLeuCysAlaGly 740  
QY 694 TTCTCTGTATGAGGAGAGCGCATGTCAGGAGATTCAGGAGCTTCACTCATCTGCGGG 753  
DB 741 TyrArgLysGlyLysLysAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysLys 760  
QY 754 AATAAGAAAGGGCGCTGCAGCTCTGGCTGGTGTGACTTCTCTGGGTTTGGGTGTGTCGA 813

DB 761 AlaLeuSerGlyArgTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArg 780  
QY 814 GCCTGGAGAAACAATGTGAGGAAAAGTCAAGGATCCCT-----GGGATCTTC 864  
DB 781 -----ProAsnTyrPheGlyValTyr 787  
QY 865 ACAGACATAGTAAGTCTTCTCTGATCCACGAA 900  
DB 788 ThrArgIleThrGlyValIleSerTrpIleGlnGln 799  
RESULT 15  
AAB24052  
ID AAB24052 standard; Protein; 802 AA.  
XX AAB24052;  
XX  
DT 25-JAN-2001 (first entry)  
XX  
DE Human PRO618 protein sequence SEQ ID NO:24.  
XX  
KW Human; tumour; diagnosis; neoplastic disease; identification; cancer;  
KW tumorigenesis; detection; neoplastic cell growth; proliferation;  
KW cytostatic; antiinflammatory; immunomodulatory; inflammatory disorder;  
KW immunological disorder.  
XX Homo sapiens.  
XX  
PN WO200053754-A1.  
XX  
XX 14-SEP-2000.  
XX  
XX 06-JAN-2000; 2000WO-US00277.  
XX  
XX 08-MAR-1999; 99WO-US05028.  
XX 12-MAR-1999; 99US-0123957.  
XX 29-MAR-1999; 99US-0126773.  
XX 21-APR-1999; 99US-0130232.  
XX 28-APR-1999; 99US-0131445.  
XX 05-OCT-1999; 99WO-US23089.  
XX 30-NOV-1999; 99WO-US28313.  
XX 02-DEC-1999; 99WO-US28551.  
XX 30-DEC-1999; 99WO-US31243.  
XX 30-DEC-1999; 99WO-US31274.  
XX  
XX (GETH ) GENENTECH INC.  
XX  
XX Baker KP, Desauvage FJ, Goddard A, Gurney AL, Klein RD, Roy MA;  
PI Wood WI;  
XX  
XX WPI; 2000-572269/53.  
XX N-PSDB; AAC58236.  
XX  
XX New isolated antibody for use in compositions and methods for the  
PT diagnosis and treatment of neoplastic cell growth and proliferation in  
PT mammals, including humans, and in monitoring tumor treatment -  
XX  
XX Claim 61; Fig 24; 195pp; English.  
XX  
XX The present invention describes an isolated antibody (Ab) that binds to  
CC one of the human proteins (P) designated PRO213, PRO1330, PRO1449,  
CC PRO237, PRO324, PRO351, PRO362, PRO615, PRO538, PRO3664, PRO618,  
CC PRO772, PRO703, PRO792 or PRO474. The Ab can be used in compositions  
CC and methods for the diagnosis and treatment of neoplastic cell growth  
CC and proliferation in mammals, including humans. Genes and polypeptides  
CC encoded by them, that are amplified in the genome of a tumour cell, can  
CC be identified and are useful targets for the treatment and prevention of  
CC certain cancers and may be used to monitor tumour treatment. Compounds  
CC that inhibit the expression or activity of the identified polypeptides  
CC can be identified and used as antagonists. Benign or malignant tumours,  
CC inflammatory disorders and immunological disorders can be treated.  
CC AAC58123 to AAC58224 represent hybridisation probes and PCR primers used



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: March 25, 2003, 07:00:26 ; Search time 51 seconds  
(without alignments)  
1930.931 Million cell updates/sec

Title: US-09-735-713A-1  
Perfect score: 1668  
Sequence: 1 atgagtctcaaatgcttat.....acatcccaactggaactaa 921

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0  
  
Searched: 221153 seqs, 53462247 residues

Total number of hits satisfying chosen parameters: 442306

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+n2p.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_spool/US09735713/runat\_18032003\_124356\_23271/app\_query.fasta\_1.1095  
-DB=PublishedApplications\_AA -QFWT=fastan -SUFFIX=rapp -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09735713 @CGN1.1.9 erunat\_18032003\_124356\_23271  
-NCPU=6 -ICPU=3 -NO\_XLPXY -NO\_WMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1626	97.5	306	9	US-09-735-713A-2
2	1608	96.4	302	9	US-09-735-713A-4
3	1356	81.3	556	10	US-09-888-615-106
4	898	53.8	164	9	US-09-735-713A-6

5	489.5	29.3	802	9	US-09-978-295A-169	Sequence 169, App
6	489.5	29.3	802	9	US-09-978-697-169	Sequence 169, App
7	489.5	29.3	802	9	US-09-978-192A-169	Sequence 169, App
8	489.5	29.3	802	9	US-09-999-832A-169	Sequence 169, App
9	489.5	29.3	802	9	US-09-978-189-169	Sequence 169, App
10	489.5	29.3	802	9	US-09-978-608A-169	Sequence 169, App
11	489.5	29.3	802	9	US-09-978-403A-169	Sequence 169, App
12	489.5	29.3	802	9	US-09-978-191A-169	Sequence 169, App
13	489.5	29.3	802	9	US-09-978-564A-169	Sequence 169, App
14	489.5	29.3	802	9	US-09-978-585A-169	Sequence 169, App
15	489.5	29.3	802	9	US-10-017-081A-169	Sequence 169, App
16	489.5	29.3	802	10	US-09-888-615-113	Sequence 113, App
17	467	28.0	453	9	US-09-978-295A-69	Sequence 69, Appl
18	467	28.0	453	9	US-09-978-697-69	Sequence 69, Appl
19	467	28.0	453	9	US-09-978-192A-69	Sequence 69, Appl
20	467	28.0	453	9	US-09-999-832A-69	Sequence 69, Appl
21	467	28.0	453	9	US-09-978-189-69	Sequence 69, Appl
22	467	28.0	453	9	US-10-174-590-64	Sequence 64, Appl
23	467	28.0	453	9	US-10-176-758-64	Sequence 64, Appl
24	467	28.0	453	9	US-10-175-737-64	Sequence 64, Appl
25	467	28.0	453	9	US-10-173-706-64	Sequence 64, Appl
26	467	28.0	453	9	US-10-175-738-64	Sequence 64, Appl
27	467	28.0	453	9	US-10-175-752-64	Sequence 64, Appl
28	467	28.0	453	9	US-10-176-482-64	Sequence 64, Appl
29	467	28.0	453	9	US-10-176-757-64	Sequence 64, Appl
30	467	28.0	453	9	US-10-176-913-64	Sequence 64, Appl
31	467	28.0	453	9	US-10-180-552-64	Sequence 64, Appl
32	467	28.0	453	9	US-10-180-557-64	Sequence 64, Appl
33	467	28.0	453	9	US-10-173-700-64	Sequence 64, Appl
34	467	28.0	453	9	US-10-174-572-64	Sequence 64, Appl
35	467	28.0	453	9	US-10-174-579-64	Sequence 64, Appl
36	467	28.0	453	9	US-10-174-582-64	Sequence 64, Appl
37	467	28.0	453	9	US-10-174-588-64	Sequence 64, Appl
38	467	28.0	453	9	US-10-175-739-64	Sequence 64, Appl
39	467	28.0	453	9	US-10-175-740-64	Sequence 64, Appl
40	467	28.0	453	9	US-10-175-743-64	Sequence 64, Appl
41	467	28.0	453	9	US-10-176-488-64	Sequence 64, Appl
42	467	28.0	453	9	US-10-176-492-64	Sequence 64, Appl
43	467	28.0	453	9	US-10-176-747-64	Sequence 64, Appl
44	467	28.0	453	9	US-10-176-750-64	Sequence 64, Appl
45	467	28.0	453	9	US-10-176-985-64	Sequence 64, Appl

ALIGNMENTS

RESULT 1  
US-09-735-713A-2  
; Sequence 2, Application US/09735713A  
; Patent No. US20020165376A1  
; GENERAL INFORMATION:  
; APPLICANT: Walke, D. Wade  
; APPLICANT: Turner, C. Alexander Jr.  
; APPLICANT: Abulin, Alejandro  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: No. US20020165376A1el Human Proteases and  
; FILE REFERENCE: LEX-0108-USA  
; CURRENT APPLICATION NUMBER: US/09/735,713A  
; CURRENT FILING DATE: 2000-12-12  
; PRIOR APPLICATION NUMBER: US 60/171,566  
; PRIOR FILING DATE: 1999-12-22  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 306  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-735-713A-2

Alignment Scores: 1.91e-152 Length: 306  
Pred. No.: 1.91e-152

Score: 1626.00 Matches: 304  
Percent Similarity: 99.35% Conservativity: 0  
Best Local Similarity: 99.35% Mismatches: 2  
Query Match: 97.48% Indels: 0  
DB: 9 Gaps: 0

US-09-735-713A-1 (1-921) x US-09-735-713A-2 (1-306)

```
Qy 1 ATGAGTCTCAAAATGCTTATTAAGCAGGAAACAAGCTGATTTTACTAGGAATAGTCTTT 60
Db 1 MetSerLeuLysMetLeuLeuSerArgAsnLysLeuLeuLeuLeuLeuGlyValPhe 20
Qy 61 TTTGAACRAGTAATCTGCARCTCTTTTCGCTCCCAAGCTCCAGTGTGGCAGAGT 120
Db 21 PheGluArgGlyLysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGlnSer 40
Qy 121 CTGGTTAAGGTACAGCTTGAATATTATTAACATTTTTCAGTCCCATCTTTGGAGGAGC 180
Db 41 LeuValLysValGlnProTrpAsnTrpPheAsnLeuPheSerArgLeuLeuGlyGlySer 60
Qy 181 CAAGTGGAGAAGGTTCTCTATCCCTGCGAGGTATCTCTGAAACAAAGGAGCAAGCATATT 240
Db 61 GlnValGluLysGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80
Qy 241 TGTGGAGAAGCATCGTCTACACAGTGGGTGATCAGCGCGGTCTACTGCATTCGAAAC 300
Db 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100
Qy 301 AGAACATTTGCTACTTTTGAATGTTACTCTGGAGAGTATGACTTAAGCCACAGAC 360
Db 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120
Qy 361 CCAGGAGAGCAACTCTCACATATTGAAACTGTCTATCATATCAGCTGCTCTCCACCAAG 420
Db 121 ProGlyGluGlnThrLeuThrIleGluThrValIleLeuHisProHisPheSerThrLys 140
Qy 421 AAACCAATGGACTATGATATTGCCCCCTTTTGAAGTGGCTGGAGCTTCCAAATTTGGCCAC 480
Db 141 LysPrometAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160
Qy 481 TTTCTGGGGCCCATATGCTTCCAGAGCTCGGGAGCAATTTGAGGCTGGTTTATTGT 540
Db 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180
Qy 541 ACAACTGAGGCTGGGGCCGCTTAACCTGAAGTGGCGTCTCTCACAAAGTCTTCAGGAA 600
Db 181 ThrThrAlaGlyTyrPglyArgLeuThrGluGlyValLeuSerGlnValLeuGlnGlu 200
Qy 601 GTGAATCTGCCTATTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTTAAACACTAAAGAGG 660
Db 201 ValAsnLeuProIleLeuThrTrpGluCysValAlaAlaLeuLeuThrLeuLysArg 220
Qy 661 CCCATCAGTGGGAAGACTTTCTTTTGACAGCTTTCCTGATGAGGAGGAGAGCGCATGT 720
Db 221 ProfilesGlyLysThrPheLeuCysThrGlyPheProaspGlyLysArgAspAlaCys 240
Qy 721 CAGGAGATTCCAGGAGTTCACTCATGTGCCGGAATTAAGAAAGGGCGCTGGACTCTGGCT 780
Db 241 GlnGlyaspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAla 260
Qy 781 GGTGTGACTTCTCGGGTTGGGCTGTGGTGGAGGCTGGAGAAACAATGTGAGGAAAGT 840
Db 261 GlyValThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnAsnValArgLysSer 280
Qy 841 GATCAAGGATCCCTGGGATCTTCACAGACATTAAGTAAAGTCTTCTCGATCCAGCAA 900
Db 281 AspGlnGlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGlu 300
Qy 901 CACATCCAAACTGGTAAAC 918
Db 301 HisIleGlnThrGlyAsn 306
```

RESULT 2

US-09-735-713A-4  
; Sequence 4, Application US/09735713A  
; Patent No. US20020165376A1  
; GENERAL INFORMATION:  
; APPLICANT: Waite, D. Wade  
; APPLICANT: Turner, C. Alexander Jr.  
; APPLICANT: Abuin, Alejandro  
; APPLICANT: Friedrich, Glenn  
; APPLICANT: Zambrowicz, Brian  
; APPLICANT: Sands, Arthur T.  
; TITLE OF INVENTION: No. US20020165376A1 Human Proteases and  
; TITLE OF INVENTION: Polynucleotides Encoding the Same  
; FILE REFERENCE: LEX-0108-USA  
; CURRENT APPLICATION NUMBER: US/09735.713A  
; CURRENT FILING DATE: 2000-12-12  
; PRIOR APPLICATION NUMBER: US 60/171,566  
; PRIOR FILING DATE: 1999-12-22  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 302  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-735-713A-4  
Alignment Scores:  
Pred. No.: 1-15e-150 Length: 302  
Score: 1608.00 Matches: 300  
Percent Similarity: 99.34% Conservativity: 0  
Best Local Similarity: 99.34% Mismatches: 2  
Query Match: 96.40% Indels: 0  
DB: 9 Gaps: 0

US-09-735-713A-1 (1-921) x US-09-735-713A-4 (1-302)

```
Qy 13 ATGTTTATAAGCAGGAAACAAGCTGATTTTACTAGGAATAGTCTTTTGAACRAGGT 72
Db 1 MetLeuIleSerArgAsnLysLeuIleLeuLeuLeuGlyIleValPhePheGluArgGly 20
Qy 73 AAATCTGCARCTCTTTCTGCTCCCAAGCTCCAGTGTGGGAGAGTCTGGTAAAGTA 132
Db 21 LysSerAlaAlaLeuSerLeuProLysAlaProSerCysGlyGlnSerLeuValLysVal 40
Qy 133 CAGCTTTGGAAATTTTAACTTTTTCAGTCGCTCTTGGAGGAAAGCAAGTGGAGAG 192
Db 41 GlnProTrpAsnTrpPheAsnIlePheSerArgIleLeuGlyGlySerGlnValGluLys 60
Qy 193 GGTTCCTATCCCTGGCAGGTATCTCTGAAACAAAGGAGAGCATATTTGTGGAGGAGC 252
Db 61 GlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIleCysGlyGlySer 80
Qy 253 ATCGTCTCACCACAGTGGGTGATCAGCGCGCTCCTGATTCATTCGCAACAGAGAAACATTTGTG 312
Db 81 IleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsnArgAsnIleVal 100
Qy 313 TCTACTTTGAAATGTTACTGCTGGAGAGTATGACTTAAGCCACAGACAGAGGAGAGCAA 372
Db 101 SerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAspProGlyGluGln 120
Qy 373 ACTCTCACTATTTGAACACTGTCATCATATCCACATTTCTCCACCAAGAAACCAATGGAC 432
Db 121 ThrLeuThrIleGluThrValIleHisProHisPheSerThrLysLysProMetAsp 140
Qy 433 TATCATATTTGCCCTTTTGAAGATGCGCTGGAGCCTTCCAATTTGCCACTTTTGTGGGCCCC 492
Db 141 TyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHisPheValGlyPro 160
Qy 493 ATATGTCTTCCAGAGCTCGGGAGCAATTTGAGGCTGGTATTATTATTGTACAACTGCAGGC 552
Db 161 IleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThrAlaGly 180
Qy 553 TGGGGCCGCTTAAGTGAAGTGGCGTCTCTCAAGTCTTTCAGAGGAGTGAATCTGCCT 612
Db 553 TGGGGCCGCTTAAGTGAAGTGGCGTCTCTCAAGTCTTTCAGAGGAGTGAATCTGCCT
```

```
Db 181 TrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnValAsnLeuPro 200
QY 613 ATTTGACCTGGAGAGTGTGGCAGCTCTTTAACACTAAGAGGCCCATCACTGGG 672
Db 201 IleLeuThrTrpGluGlyValAlaLeuLeuThrLeuLysArgProIleSerGly 220
QY 673 AAGACCTTTCTTTGACAGGTTTCTGATGGAGGAGAGCGCATGTTCAGGAGATTCA 732
Db 221 LysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCysGlnGlyAspSer 240
QY 733 GGAGGTTCATCTGTCGCCGGAATAAGAAAGGGCCCTGGACTCTGCTGGTGTGACTTCC 792
Db 241 GlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAlaGlyValThrSer 260
QY 793 TGGGGTTTGGCTGTGGTGGAGCTGGAGAAACAATGTGAGGAAAAGTGCATCAAGGATCC 852
Db 261 TrpGlyLeuGlyCysGlyArgGlyTrpArgAsnAsnValArgLysSerAspGlnGlySer 280
QY 853 CCTGGGATCTTCACAGACATTAGTAAAGTCTTCTCGGATCCACGAAACATCCAAACT 912
Db 281 ProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIleGlnThr 300
QY 913 GGTAAC 918
Db 301 GlyAsn 302
RESULT 3
US-09-888-615-106
; Sequence 106, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY J.
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEL, SEAN.
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; PRIOR FILING DATE: 2001-06-26
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 106
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-106
```

```
Alignment Scores:
Pred. No.: 1,19e-125 Length: 556
Score: 1356.00 Matches: 262
Percent Similarity: 87.9% Conservative: 29
Best Local Similarity: 85.34% Mismatches: 8
Query Match: 81.29% Indels: 8
DB: 10 Gaps: 2
```

US-09-735-713A-1 (1-921) x US-09-888-615-106 (1-556)

```
QY 1 ATGAGTCTCAAAATGCTTATAAGCAGGAAACAAGCTGATTTTACTAGGAATAGTCTTT 60
Db 1 MetSerLeuLysMetLeuIleSerArgAsnLysLeuIleLeuLeuGlyIleValPhe 20
QY 61 TTGTGAACRAGGTAATCTGCARCTCTTTCGCTCCCAAGAGTCCCAAGTGTGGGAGAGT 120
Db 21 PheGluArgGlyLysSerAlaThrLeuSerLeuProLysAlaProSerCysGlyGlnSer 40
QY 121 CTGGTTAAGGTACAGCCTTGGAAATTTATTTACATTTTTCAGTCGCATTTCTTGAGGAAGC 180
Db 41 LeuValLysValGlnProTrpAsnTyPheAsnIlePheSerArgIleLeuGlyGlySer 60
```

```
QY 181 CAAGTGGAAGAGGTTTCCTATCCCTGGCAGGTATCTCTGAACAAAGGAGCAGATATT 240
Db 61 GlnValGluGlySerTyrProTrpGlnValSerLeuLysGlnArgGlnLysHisIle 80
QY 241 TGTGGAGGAAGCATCGTCTCACACAGTGGGTGATCACGGCGGCTCACATGCAATTGCAAC 300
Db 81 CysGlyGlySerIleValSerProGlnTrpValIleThrAlaAlaHisCysIleAlaAsn 100
QY 301 AGAACATTTGCTCTACTTTGAATCTTACTGTGGAGAGTATGACTTAAAGCCAGACAGAC 360
Db 101 ArgAsnIleValSerThrLeuAsnValThrAlaGlyGluTyrAspLeuSerGlnThrAsp 120
QY 361 CCAGGAGAGCAAACTCTCACTATTGAAACTGTCATCATACATCCACATTTCTCCACCAAG 420
Db 121 ProGlyGluGlnThrLeuThrIleGluThrValIleIleHisProHisPheSerThrLys 140
QY 421 AAACCAATGGACTATGATATTGCCCTTTTGAAGATGGCTGGAGCCCTTCCAATTTGGCCAC 480
Db 141 LysProMetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHis 160
QY 481 TTTGTGGGCCCATATGCTCTCCAGAGCTGCGGAGCAATTTGAGGCTGCTTTTATTTCGT 540
Db 161 PheValGlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCys 180
QY 541 ACAACTGCAGGCTGGGGCCGCTTAAGGTGGGCTCTCTCTCAAAAGTCTTTCAGGAA 600
Db 181 ThrThrAlaGlyTrpGlyArgLeuThrGluGlyGlyValLeuSerGlnValLeuGlnGlu 200
QY 601 GTGAATCTCCCTATTTTGACCTGGGGAAGAGTGTGTGGCAGCTGTGTAACTAACAGAGG 660
Db 201 ValAsnLeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArg 220
QY 661 CCCATCAGTGGGAAGACCTTTCTTCCACAGGTTTTTCTGATGGAGGAGACGCATCT 720
Db 221 ProIleSerGlyLysThrPheLeuCysThrGlyPheProAspGlyGlyArgAspAlaCys 240
QY 721 CAGGAGATTCAGAGGTTTCACTCATGTGCCGGAATAAGAAAGGCGCTGGACTCTGGCT 780
Db 241 GlnGlyAspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpAsp---Ser 259
QY 781 GGTGTGACTTCTCGGGTTTGGGCTGTGTCGAGGCTGGAGAAACAACTGTGAGGAAAGT 840
Db 260 GlyTrpSerIleTrpGluAlaGlnValGlySerLeuGluSerArgSerSerArgPro 279
QY 841 GATCAAGGATCCCT-----GGGATCTTTCACAGACATTTAGTAAA 879
Db 280 SerLeuGlyAsnLysValArgLeuCysLeuThrAsnAsnPhePheLysLysLeuAlaGly 299
QY 880 GTGCTTTCTCGGATCCACGAA 900
Db 300 CysGlyThrTrpCysSerGlu 306
RESULT 4
US-09-735-713A-6
; Sequence 6, Application US/09735713A
; Patent No. US20020165376A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Abuin, Alejandro
; APPLICANT: Friedrich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020165376A1el Human Proteases and
; FILE OF INVENTION: Polynucleotides Encoding the Same
; FILE REFERENCE: LEX-0108-USA
; CURRENT APPLICATION NUMBER: US/09/735,713A
; CURRENT FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: US 60/171,566
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
```

```

; LENGTH: 164
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-735-713A-6

Alignment Scores:
Pred. No.: 1.53e-80 Length: 164
Score: 898.00 Matches: 164
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 53.84% Indels: 0
DB: 9 Gaps: 0

US-09-735-713A-1 (1-921) x US-09-735-713A-6 (1-164)
QY 427 ATGCATGATGATATGCCCTTTTCAAGATGCTGGAGCCCTCCAAATTTGCCACTTTGTG 486
Db 1 MetAspTyrAspIleAlaLeuLeuLysMetAlaGlyAlaPheGlnPheGlyHisPheVal 20
QY 487 GGGCCCATATCTCTCCAGAGCTCGGGAGCAATTTGAGGCTGGTTTATTGTACAACT 546
Db 21 GlyProIleCysLeuProGluLeuArgGluGlnPheGluAlaGlyPheIleCysThrThr 40
QY 547 GCAGCTGGGGCCCTTAACCTGAAGTGGCGTCTCTCACAAAGTCTTCAGGAAAGTGAAT 606
Db 41 AlaGlyTrpGlyArgLeuThrGluGlyCysValLeuSerGlnValLeuGlnGluValAsn 60
QY 607 CTGCTATTTTGACCTGGGAAGAGTGTGTGGCAGCTCTGTAAACACAAAGAGGCCCATC 666
Db 61 LeuProIleLeuThrTrpGluGluCysValAlaAlaLeuLeuThrLeuLysArgProIle 80
QY 667 AGTGGGAAGACTTTCTTTGACAGAGTTTCTCTGATGGAGGAGAGCATGTACGGGA 726
Db 81 SerGlyLysThrPheLeuCysThrGlyPheProaspGlyArgaspAlaCysGlnGly 100
QY 727 GATTGAGAGGTTCACTCATGTGCCGAATAAGAAAGGGCCCTGGACTCTGGCTGGTGTG 786
Db 101 AspSerGlyGlySerLeuMetCysArgAsnLysLysGlyAlaTrpThrLeuAlaGlyVal 120
QY 787 ACTTCCCTGGGTTGGCTGTGTCGAGGTCGGAGAACATATGTGAGAAAGTATGATCAA 846
Db 121 ThrSerTrpGlyLeuGlyCysGlyArgGlyTrpArgAsnValArgLysSerAspGln 140
QY 847 GGATCCCTGGGATCTTCACAGACATAGTAAAGTGTCTTCCTGGATCCACGACACATC 906
Db 141 GlySerProGlyIlePheThrAspIleSerLysValLeuSerTrpIleHisGluHisIle 160
QY 907 CAAACTGGTATAC 918
Db 161 GlnThrGlyAsn 164

RESULT 5
US-09-978-295A-169
; Sequence 169, Application US/09978295A
; Patent No. US20020156006A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
```

60/080328	PRIOR APPLICATION NUMBER: 60/080328
60/080329	PRIOR FILING DATE: 1998-04-01
60/080330	PRIOR APPLICATION NUMBER: 60/080330
60/080331	PRIOR FILING DATE: 1998-04-01
60/080332	PRIOR APPLICATION NUMBER: 60/080332
60/080333	PRIOR FILING DATE: 1998-04-01
60/080334	PRIOR APPLICATION NUMBER: 60/080334
60/080335	PRIOR FILING DATE: 1998-04-01
60/081070	PRIOR APPLICATION NUMBER: 60/081070
60/081071	PRIOR FILING DATE: 1998-04-08
60/081072	PRIOR APPLICATION NUMBER: 60/081072
60/081073	PRIOR FILING DATE: 1998-04-08
60/081074	PRIOR APPLICATION NUMBER: 60/081074
60/081075	PRIOR FILING DATE: 1998-04-08
60/081195	PRIOR APPLICATION NUMBER: 60/081195
60/081196	PRIOR FILING DATE: 1998-04-08
60/081203	PRIOR APPLICATION NUMBER: 60/081203
60/081204	PRIOR FILING DATE: 1998-04-09
60/081229	PRIOR APPLICATION NUMBER: 60/081229
60/081230	PRIOR FILING DATE: 1998-04-09
60/081955	PRIOR APPLICATION NUMBER: 60/081955
60/081956	PRIOR FILING DATE: 1998-04-15
60/081817	PRIOR APPLICATION NUMBER: 60/081817
60/081818	PRIOR FILING DATE: 1998-04-15
60/081819	PRIOR APPLICATION NUMBER: 60/081819
60/082568	PRIOR FILING DATE: 1998-04-15
60/082569	PRIOR APPLICATION NUMBER: 60/082569
60/082570	PRIOR FILING DATE: 1998-04-21
60/082700	PRIOR APPLICATION NUMBER: 60/082700
60/082701	PRIOR FILING DATE: 1998-04-22
60/082797	PRIOR APPLICATION NUMBER: 60/082797
60/082798	PRIOR FILING DATE: 1998-04-22
60/082796	PRIOR APPLICATION NUMBER: 60/082796
60/082797	PRIOR FILING DATE: 1998-04-23
60/083336	PRIOR APPLICATION NUMBER: 60/083336
60/083337	PRIOR FILING DATE: 1998-04-27
60/083322	PRIOR APPLICATION NUMBER: 60/083322
60/083323	PRIOR FILING DATE: 1998-04-28
60/083392	PRIOR APPLICATION NUMBER: 60/083392
60/083495	PRIOR FILING DATE: 1998-04-29
60/083496	PRIOR APPLICATION NUMBER: 60/083495
60/083497	PRIOR FILING DATE: 1998-04-29
60/083498	PRIOR APPLICATION NUMBER: 60/083499
60/083545	PRIOR FILING DATE: 1998-04-29
60/083546	PRIOR APPLICATION NUMBER: 60/083545
60/083547	PRIOR FILING DATE: 1998-04-29
60/083554	PRIOR APPLICATION NUMBER: 60/083554
60/083555	PRIOR FILING DATE: 1998-04-29
60/083556	PRIOR APPLICATION NUMBER: 60/083558
60/083557	PRIOR FILING DATE: 1998-04-29
60/083558	PRIOR APPLICATION NUMBER: 60/083559
60/083559	PRIOR FILING DATE: 1998-04-29
60/083560	PRIOR APPLICATION NUMBER: 60/083560
60/083561	PRIOR FILING DATE: 1998-04-29
60/083562	PRIOR APPLICATION NUMBER: 60/083563
60/083564	PRIOR FILING DATE: 1998-04-29
60/083565	PRIOR APPLICATION NUMBER: 60/083566
60/083567	PRIOR FILING DATE: 1998-05-05
60/084414	PRIOR APPLICATION NUMBER: 60/084414
60/084415	PRIOR FILING DATE: 1998-05-06
60/084416	PRIOR APPLICATION NUMBER: 60/084416
60/084417	PRIOR FILING DATE: 1998-05-06
60/084637	PRIOR APPLICATION NUMBER: 60/084637

?	PRIOR FILING DATE:	1998-05-07	
?	PRIOR APPLICATION NUMBER:	60/084639	
?	PRIOR FILING DATE:	1998-05-07	
?	PRIOR APPLICATION NUMBER:	60/084640	
?	PRIOR FILING DATE:	1998-05-07	
?	PRIOR APPLICATION NUMBER:	60/084598	
?	PRIOR FILING DATE:	1998-05-07	
?	PRIOR APPLICATION NUMBER:	60/084600	
?	PRIOR FILING DATE:	1998-05-07	
?	PRIOR APPLICATION NUMBER:	60/084627	
?	PRIOR FILING DATE:	1998-05-07	
?	PRIOR APPLICATION NUMBER:	60/084643	
?	PRIOR FILING DATE:	1998-05-07	
?	PRIOR APPLICATION NUMBER:	60/085339	
?	PRIOR FILING DATE:	1998-05-13	
?	PRIOR APPLICATION NUMBER:	60/085338	
?	PRIOR FILING DATE:	1998-05-13	
?	PRIOR APPLICATION NUMBER:	60/085323	
?	PRIOR FILING DATE:	1998-05-13	
?	PRIOR APPLICATION NUMBER:	60/085582	
?	PRIOR FILING DATE:	1998-05-15	
?	PRIOR APPLICATION NUMBER:	60/085700	
?	PRIOR FILING DATE:	1998-05-15	
?	PRIOR APPLICATION NUMBER:	60/085689	
?	PRIOR FILING DATE:	1998-05-15	
?	PRIOR APPLICATION NUMBER:	60/085579	
?	PRIOR FILING DATE:	1998-05-15	
?	PRIOR APPLICATION NUMBER:	60/085580	
?	PRIOR FILING DATE:	1998-05-15	
?	PRIOR APPLICATION NUMBER:	60/085573	
?	PRIOR FILING DATE:	1998-05-15	
?	PRIOR APPLICATION NUMBER:	60/085704	
?	PRIOR FILING DATE:	1998-05-15	
?	PRIOR APPLICATION NUMBER:	60/085697	

Alignment Scores:

Alignment Scores:		
Pred. No.:	6.72e-40	Length:
Score:	489.50	Matches:
Percent Similarity:	56.75%	Conservative:
Best Local Similarity:	39.29%	Mismatches:
Query Match:	29.35%	Indels:
DB:	9	Gaps:
		802
		99
		44
		86
		23
		5

US-09-735-713A-1 (1-921) x US-09-978-295A-169 (1-802)

Qy	160	AGTCGCATCTTGGAGGAAGCCCAAGTGAGAAAGGGTTCCCTATACCTCGCAGGATGATCTCTGT      :::     :::     :::     :::     :::     :::     :::     :::     :::
Db	566	SerArgIleValGlyAlaValSerSerGluGlycLutrpProItrpGlnAlaSerLeu      :::     :::     :::     :::     :::     :::     :::     :::
Qy	220	AAACAAGAGCAAGATATTTGTGGAGGAAGCATCGTCTCACCAACAGTAGGGGTGATCACG :::     ::::   :::     :::     :::     :::     :::     :::     :::
Db	586	GlnValArgGlyArgHisIleCysGlyAlaLeuIleAlaAspArgTrpValIleThr :::     ::::   :::     :::     :::     :::     :::     :::     :::
Qy	280	CGCGCTCAGTCGATTGCAACACAGAACAATGTGTCTACTTTG-----AATGTACTTGCT      :::     :::     :::     :::     :::     :::     :::     :::
Db	606	AlaAlaHisCysPheGlnGluAspSerMetAlaSerThrValLeuTrpThrValPheLeu      :::     :::     :::     :::     :::     :::     :::     :::
Qy	334	GGAGAGTATGACTTAAGCCACAGACACCAGGAGAGCAAACCTCTCACTATTGAACTGTC      :::     :::     :::     :::     :::     :::     :::     :::
Db	626	GlyLysValTrpGlnAsnSerArgTrpProGlycLutrpValSerPheLysValSerArgLeu      :::     :::     :::     :::     :::     :::     :::     :::
Qy	394	ATCATACATCCACATTTCCTCCACCAAGAACCAATGGACTATCATATTGGCCCTTTTGAAG :::     :::     :::     :::     :::     :::     :::     :::
Db	646	LeuLeuHisPro---TyrHisGluUspSerHisAspTyrAspValAlaLeuLeuGln :::     :::     :::     :::     :::     :::     :::     :::
Qy	454	ATGCGCTGGAGCCTTCCAATTGGCCACTTTGTGGGGCCCATAATGTCTTCCAGAGCTGCCGG :::        :::     :::     :::     :::     :::     :::     :::
Db	665	LeuAspHisProValValArgSerAlaAlaValArgProValCysLeuProAlaArgSer :::        :::     :::     :::     :::     :::     :::     :::
Qy	514	GAGCAATTGGAGCGCTTTTATTGTACAACCTCCAGGCTGGGGCCGCTTAACTGAAGGT              :::     :::     :::     :::     :::     :::     :::
Db	685	HisPheGluProGlyLeuHisCysTrpIleThrGlyTrpGlyAlaLeuArgGluGly              :::     :::     :::     :::     :::     :::     :::



; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-5-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15

; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:
Pred. No.: 6,72e-40 Length: 802
Score: 489.50 Matches: 99
Percent Similarity: 56.75% Conservatives: 44
Best Local Similarity: 39.29% Mismatches: 86
Query Match: 29.35% Indels: 23
DB: 9 Gaps: 5

US-09-735-713a-1 (1-921) x US-09-978-697-169 (1-802)

QY 160 AGTCGCATTCTTGGAGGAAGCCAAAGTGGAGAGGGTTCCTATCCCTGGCAGGTATCTCTG 219
|||||
Db 566 SerArgIleValGlyGlyAlaValSerSerGluGlyGluTrpProTrpGlnAlaSerLeu 585
QY 220 AAACAAGGCGAGGAGCATATTTGCGAGGAGGAGCATGCTCACCACAGTGGGTGATCAGC 279
::: ||| ::|||
Db 586 GlnValArgGlyArgHisIleCysGlyAlaLeuIleAlaAspArgTrpValIleThr 605
QY 280 GCGGCTCACTGCATTGCAACACAAACATTTGCTCTACTTTG-----AATGTTACTGCT 333
|||||
Db 606 AlaAlaHisCysPheGlnGluAspSerMetAlaSerThrValLeuTrpThrValPheLeu 625
QY 334 GGAGGATATGACTTAAGCCAGACAGACCCAGAGGAGCAAACTCTCATTATTTAAACTGTC 393
|||||
Db 626 GlyLysValTrpGlnAsnSerArgTrpProGlyGluValSerPheLysValSerArgLeu 645
QY 394 ATCATATCCACATTTCTCCACCAAGAACCAATGAGCATATATGATATTTGCCCTTTGAAG 453
::: ||| ::|||
Db 646 LeuLeuHisPro---TyrHisGluGluAspSerHisAspTyrAspValAlaLeuLeuGln 664
QY 454 ATGGCTGGAGCCCTTCCAAATTTGCCACTTTGTGGGGCCCATATGCTCTCCAGAGTCGCG 513
::: |||
Db 665 LeuAspHisProValValArgSerAlaAlaValArgProValCysLeuProAlaArgSer 684
QY 514 GAGCAATTTGAGGCTGGTTTTATTGTACAACTGCAGGCTGGGGCCGCTTTAACTGAAGT 573
|||||
Db 685 HisPhePheGluProGlyLeuHisCysTrpIleThrGlyTrpGlyAlaLeuArgGluGly 704
QY 574 GCGCTCTCTCACAAGTCTTGCAGGAGTGAATCTGCCTATTGCTGACCTGGGAGAGTGT 633
||| ::|||
Db 705 GlyProIleSerAsnAlaLeuGlnLysValAspValGlnLeuIleProGlnAspLeuCys 724
QY 634 GTGGCAGCTCTGTAAACACTAAAGAGGCCCATCAGTGGGAAGACCTTTCTTTCACAGGT 693
|||
Db 725 SerGluAlaTyrArgTyrGlnValThrPro-----ArgMetLeuCysAlaGly 740
QY 694 TTTCTCTGATGGAGGAGAGACGATGTCAGGAGATTTCAGGAGGTTCACTCATGTGCGCG 753
::: ||| ::|||
Db 741 TyrArgLysGlyLysLysAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysLys 760
QY 754 AATAAGAAAGGGCCCTGGACTCTGGCTGGTGTGACTTCTCGGGGTTGGGCTGTGGTCGA 813
||| |||
Db 761 AlaLeuSerGlyArgTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArg 780
QY 814 GGCTGGAGAAACAATGTGAGGAAAAGTATCAAGGATCCCT-----GGATCTTC 864
|||
Db 781 -----ProAsnTyrPheGlyValTyr 787
QY 865 ACAGACATTTAGTAAAGTCTTTCTCTGATCCAGCAA 900
||| |||
Db 788 ThrArgIleThrGlyValIleSerTrpIleGlnGln 799

RESULT 7
US-09-978-192A-169
; Sequence 169, Application US/09978192A

Patent No. US2002017753A1  
GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Baker, Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kijavits, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630P1C9  
CURRENT APPLICATION NUMBER: US/09/978,192A  
CURRENT FILING DATE: 2001-10-15  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29



APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630P1C63  
CURRENT APPLICATION NUMBER: US/09/999,832A  
CURRENT FILING DATE: 2001-10-24  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084639  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084598  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084600







```
Db 705 GlyProIleSerAsnAlaLeuGlnLysValAspValGlnLeuIleProGlnAspLeuCys 724      |||  ::::|  |||||::|::|::|::|  ::::|  |||
Qy 634 GTGCACGCTGTTTAAACACATAAGAGCCCATCAGTGGGAAGACCITTTCTTTCACACAGGT 693      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 725 SerGluAlaTyrArgTyrGlnValThrPro-----ArgMetLeuCysAlaGly 740      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy 694 TTTCCTGATGGAGGAGAGACGATGTCAGGAGATTCAGGAGTTTCACATCATGTGCCGG 753      ::  |||  ::|||::|::|::|::|::|::|  |||::|::|::|::|::|::|::|
Db 741 TyrArgLysGlyLysAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysLys 760      ::  |||  ::|||::|::|::|::|::|::|  |||::|::|::|::|::|::|::|
Qy 754 AATAAGAAAGGGCGCTGCATCTGGCTGGTGTGACTTCCTGGGGTTTGGGCTGTGGTCGA 813      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 761 AlaLeuSerGlyArgTyrPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArg 780      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy 814 GGCTGGAGAAACAATGTGAGAAAGTGATCAAGGATCCCTC-----GGGATCTTC 864      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 781 -----ProAsnTyrPheGlyValTyr 787      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Qy 865 ACAGACATTAGTAAAGTGCTTCTCGATCCACGAA 900      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 788 ThrArgIleThrGlyValIleSerTrpIleGlnGln 799      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

RESULT 11
US-09-978-191A-169
; Sequence 169, Application US/09978191A
; Publication No. US20030050239A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C4
; CURRENT APPLICATION NUMBER: US/09/978,191A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918595
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
```

```
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
```

## RESULT 12

GENERAL INFORMATION:  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Baker Kevin P.  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan  
APPLICANT: Ferrara, Napoleon  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, J. Christopher  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth J.  
APPLICANT: Kljavin, Ivar J.  
APPLICANT: Kuo, Sophia S.  
APPLICANT: Napier, Mary A.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Shelton, David L.  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
TITLE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: P2630PIC17  
CURRENT APPLICATION NUMBER: US/09/978,403A  
CURRENT FILING DATE: 2002-03-19  
PRIOR APPLICATION NUMBER: 09/918585  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545

```

; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:
Pred. No.: 6,72e-40 Length: 802
Score: 489.50 Matches: 99
Percent Similarity: 56.75% Conservative: 44
Best Local Similarity: 39.29% Mismatches: 86
Query Match: 29.35% Indels: 23
DB: 9 Gaps: 5

US-09-735-713a-1 (1-921) x US-09-978-403A-169 (1-802)
Qy 160 AGTCGATCTTTGGAGGAGCAAGTGGAGAGGGTTCCTATCCCTGGCAGGATATCTGTG 219
Db 566 SerArgileValGlyAlaValSerSerGluGlyGluTrpProTrpGlnAlaSerLeu 585
Qy 220 AAACAAGGCAAGCATATTTGTGGAGCAAGCATGCTCCACACAGTGGGTGATCAGG 279
Db 586 GlnValArgGlyArgHisIleGlyAlaLeuAlaAspArgTrpValIleThr 605
Qy 280 GCGGCTCACTGCATGTCGCAACACATTTGTGTACTTTTG-----AATGTTACTGCT 333
Db 586 SerArgileValGlyAlaValSerSerGluGlyGluTrpProTrpGlnAlaSerLeu 585
```

```

Db 606 AlaAlaHisCysPheGlnGluAspSerMetalaSerThrValLeuTrpThrValPheLeu 625
Qy 334 GGAGAGTATGACTTAAGCCAGACAGACCAGAGAGCAAACTCTCACTATTGAACTGTC 393
Db 626 GlyLysValTrpGlnAsnSerArgTrpProGlyGluValSerPheLysValSerArgLeu 645
Qy 394 ATCATACATCCCATTTCTCCACCAAGAACCAATGGCATATGATATGATATGATATGAT 453
Db 646 LeuLeuHisPro---TyrHisGluGluAspSerHisAspTyrAspValAlaLeuGln 664
Qy 454 ATGGCTGGAGCCTTCCAAATTTGGCCACTTTGGGGCCCATATGCTTCCAGAGTGGCG 513
Db 665 LeuAspHisProValValArgSerAlaAlaValArgProValCysLeuProAlaArgSer 684
Qy 514 GAGCAATTTGAGGCTGGTTTATTTTGTACAACTGCAGGCTGGGGCCGCTTAAGTGT 573
Db 685 HisPhePheGluProGlyLeuHisCysTrpIleThrGlyTrpGlyAlaLeuArgGluGly 704
Qy 574 GCGCTCTCTCAAGTCTTGCAGGAAGTGAATCTGCCTTATTTTGCCTGGGAAGAGTGT 633
Db 705 GlyProIleSerAsnAlaLeuGlnLysValAspValGlnLeuIleProGlnAspLeuCys 724
Qy 634 GTGCAGCTCTGTTAACTAAAGAGGCCCATCACTGAGGAGCCCTTTCTTTCCACAGCT 693
Db 725 SerGluAlaTyrArgTyrGlnValThrPro-----ArgMetLeuCysAlaGly 740
Qy 694 TTTCTGTATGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 753
Db 741 TyrArgLysGlyLysAlaCysGlnGlyAspSerGlyGlyProLeuValCysLys 760
Qy 754 AATAAGAAAGGGGCTGCTCTGCTGTGTGACTTCTCTGGGTTTGGGCTGTGTGTCA 813
Db 761 AlaLeuSerGlyArgTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArg 780
Qy 814 GCGTGGAGAAACAATGTGAGGAGAAAGTATCAAGGATCCCT-----GGGATCTTC 864
Db 781 -----ProAsnTyrPheGlyValTyr 787
Qy 865 ACAGACATTAAGTAAAGTCTTTCTTCTGGATCCACGAA 900
Db 788 ThrArgIleThrGlyValIleSerTrpIleGlnGln 799

RESULT 13
US-09-978-564A-169
; Sequence 169, Application US/09978564A
; Publication No. US20030050241A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
```

APPLICANT: Wood, William I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE REFERENCE: P2630P1C25  
CURRENT APPLICATION NUMBER: US/09/978,564A  
CURRENT FILING DATE: 2001-10-16  
PRIOR APPLICATION NUMBER: 09/918595  
PRIOR FILING DATE: 2001-07-30  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064249  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/065311  
PRIOR FILING DATE: 1997-11-13  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/077450  
PRIOR FILING DATE: 1998-03-10  
PRIOR APPLICATION NUMBER: 60/077632  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077641  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077649  
PRIOR FILING DATE: 1998-03-11  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078004  
PRIOR FILING DATE: 1998-03-13  
PRIOR APPLICATION NUMBER: 60/078886  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078936  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/078939  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079656  
PRIOR FILING DATE: 1998-03-26  
PRIOR APPLICATION NUMBER: 60/079664  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079689  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079786  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/079920  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/079923  
PRIOR FILING DATE: 1998-03-30  
PRIOR APPLICATION NUMBER: 60/080105  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080107  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080194  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/080327  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080328  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080333  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/080334  
PRIOR FILING DATE: 1998-04-01  
PRIOR APPLICATION NUMBER: 60/081070  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081049

PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081071  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081195  
PRIOR FILING DATE: 1998-04-08  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081955  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081819  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081952  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081838  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082568  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082569  
PRIOR FILING DATE: 1998-04-21  
PRIOR APPLICATION NUMBER: 60/082704  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082804  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082700  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082797  
PRIOR FILING DATE: 1998-04-22  
PRIOR APPLICATION NUMBER: 60/082796  
PRIOR FILING DATE: 1998-04-23  
PRIOR APPLICATION NUMBER: 60/083336  
PRIOR FILING DATE: 1998-04-27  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083392  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083495  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083496  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083499  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083554  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083558  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083559  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083500  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/083742  
PRIOR FILING DATE: 1998-04-30  
PRIOR APPLICATION NUMBER: 60/084366  
PRIOR FILING DATE: 1998-05-05  
PRIOR APPLICATION NUMBER: 60/084414  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084441  
PRIOR FILING DATE: 1998-05-06  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084639  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084640  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084598  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07

;; PRIOR APPLICATION NUMBER: 60/084627  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/084643  
;; PRIOR FILING DATE: 1998-05-07  
;; PRIOR APPLICATION NUMBER: 60/085339  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085338  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085323  
;; PRIOR FILING DATE: 1998-05-13  
;; PRIOR APPLICATION NUMBER: 60/085582  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085700  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085689  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085579  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085580  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085573  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085704  
;; PRIOR FILING DATE: 1998-05-15  
;; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:  
Pred. No.: 6,72e-40 Length: 802  
Score: 489.50 Matches: 99  
Percent Similarity: 56.75% Conservativity: 44  
Best Local Similarity: 39.29% Mismatches: 86  
Query Match: 29.35% Indels: 23  
DB: 9 Gaps: 5

US-09-735-713a-1 (1-921) x US-09-978-564A-169 (1-802)

QY 160 AGTCGATCTTGGAGAACCAAGTGGAGAGGGTCTCTATCCCTGGCAGGTATCTCTG 219  
DB 566 SerArgileValGlyAlaValSerSerGlyGluTrpProTrpGlnAlaSerLeu 585  
QY 220 AACAAAGCAGAGCATATTTGTGAGGAGCATCTCTCACCAGCTGGGTGATCAG 279  
DB 586 GlnValArgGlyArgHisIleCysGlyAlaLeuIleAlaAspArgTrpValIleThr 605  
QY 280 GCGGCTCACTGCTCCAAACAGACATTTGCTCTACTTTG-----AATGTTACTGCT 333  
DB 606 AlaAlaHisCysPheGlnGluAspSerMetAlaSerThrValLeuTrpThrValPheLeu 625  
QY 334 GGAGAGTATGACTTAAGCCAGACAGACCAGGAGCAAACTCTCACTATTGAACTGTC 393  
DB 626 GlyLysValTrpGlnAsnSerArgTrpProGlyGluValSerPheLysValSerArgLeu 645  
QY 394 ATCATATCCACATTTCTCCACCAGAACCAATCGACTATGATATTCCTTTTGAAG 453  
DB 646 LeuLeuHisPro---TyrHisGluGluAspSerHisAspTyrAspValAlaLeuGln 664  
QY 454 ATGCTGGAGCCCTCCAAATTTGGCCACTTTGTGGGCCCATATGCTTCCAGAGCTCGG 513  
DB 665 LeuAspHisProValValValArgSerAlaAlaValArgProValCysLeuProAlaArgSer 684  
QY 514 GACAAATTTGAGCTGGTTTATTTTGTACAACTGCAGGCTGGGGCCGCTTAAGTGAAGT 573  
DB 685 HisPhePheGluProGlyLeuHisCysTrpIleThrGlyTrpGlyAlaLeuArgGluGly 704  
QY 574 GCGCTCTCACAAGTCTTCGAGGAGTGAATCTGCTCTATTTGACCTGGGGAAGAGTGT 633  
DB 705 GlyProIleSerAsnAlaLeuGlnLysValAspValGlnLeuIleProGlnAspLeuCys 724  
QY 634 GTGGCAGCTCTGTTAACTAAAGAGGCCCATCATCAGTGGGAGACCTTTCTTTCACAGGT 693  
DB 725 SerGluAlaTrpArgTyrGlnValThrPro-----ArgMetLeuCysAlaGly 740  
QY 694 TTTCTCTGATGGAGGAGACGCATCTCAGGGAGATTCAGGAGGTTCACCTCATGTGCCGG 753

DB 741 TyrArgLysGlyLysAspAlaCysGlnGlyAspSerGlyGlyProLeuValCysLys 760  
QY 754 AATAAGAAAGGGCGCTGGACTCTGGCTGTGTGACTTCTCTGGGGTGGGCTGTGGTCA 813  
DB 761 AlaLeuSerGlyArgTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArg 780  
QY 814 GCGTGGAGAAACAATGTGAGGAAAGTGATCAAGGATCCCT-----GGGATCTTC 864  
DB 781 -----ProAsnTyrPheGlyValTyr 787  
QY 865 ACAGACATTAGTAAAGTCTTCTCTGGATCCACGAA 900  
DB 788 ThrArgIleThrGlyValIleSerTrpIleGlnGln 799

## RESULT 14

US-09-978-585A-169  
; Sequence 169, Application US/09978585A  
; Publication No. US20030049633A1

## GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi  
; APPLICANT: Baker Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan  
; APPLICANT: Ferrara, Napoleon  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Kuo, Sophia S.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James.  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Shelton, David L.  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; FILE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: P2630PIC15  
; CURRENT APPLICATION NUMBER: US/09/978,585A  
; CURRENT FILING DATE: 2001-10-16  
; NUMBER OF SEQ ID NOS: 624  
; Prior Application removed - See File Wrapper or Palm  
; SEQ ID NO 169  
; LENGTH: 802  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-978-585A-169

Alignment Scores:  
Pred. No.: 6,72e-40 Length: 802  
Score: 489.50 Matches: 99  
Percent Similarity: 56.75% Conservativity: 44  
Best Local Similarity: 39.29% Mismatches: 86  
Query Match: 29.35% Indels: 23  
DB: 9 Gaps: 5

US-09-735-713a-1 (1-921) x US-09-978-585A-169 (1-802)

QY 160 AGTCGATCTTGGAGAACCAAGTGGAGAGGGTCTCTATCCCTGGCAGGTATCTCTG 219  
DB 566 SerArgileValGlyAlaValSerSerGlyGluTrpProTrpGlnAlaSerLeu 585



Db 761 AlaLeuSerGlyArgTrpPheLeuAlaGlyLeuValSerTrpGlyLeuGlyCysGlyArg 780  
QY 814 GGCTGGAGAAACAATGTGAGGAAAGTCAAGGATCCCT-----GGGATCTTC 864  
Db 781 -----ProAsnTyrPheGlyValTyr 787  
QY 865 ACAGACATTAGTAAAGTGTCTTCTGGATCCACGAA 900  
Db 788 ThrArgIleThrGlyValIleSerTrpIleGlnGln 799

Search completed: March 25, 2003, 07:24:21  
Job time : 56 secs

**THIS PAGE BLANK (USPTO)**